

STIC-Biot ch/ChemLib

82001

From: Schnizer, Holly  
Sent: Tuesday, December 10, 2002 9:30 AM  
To: STIC-Biotech/ChemLib  
Subject: seq. search for appl. no. 09/444,281

RECEIVED

DEC 10 2002

STIC-Biotech/ChemLib  
(STIC)

Please search the commercial and interference databases for:

SEQ ID NOs: 35 and 36 and,

a peptide with the following sequence (SEQ ID NO: 35)(SEQ ID NO:27)(SEQ ID NO:35)

Thank you.

Holly Schnizer  
AU 1653  
CM1-9E09  
305-3722  
mailbox: CM1-9B01

9E09

Searcher: Point of Contact  
Phone: P. Sheppard  
Location: Telephone number: (703) 308-4499  
Date Picked Up: \_\_\_\_\_  
Date Completed: 12/13/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

QY	4	KWPWPMPWRKRRHLEPFEALPIMLLKWPWPW	34
		:       :	
Db	1185	RMQWWSNPRRGRC-----CWQWWSW	1204

## RESULT 2

09JH31 ID 09JH31 PRELIMINARY; PRT; 746 AA.  
 AC 09JH31;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE ORF1.  
 OS TT virus.  
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.  
 OX NCBI\_TaxID=68887;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TJN02;  
 RA Okamoto H.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TJN02;  
 RA MEDLINE=20456801; PubMed=11003468;  
 RA Ukita M., Okamoto H., Nishizawa T., Tawara A., Takahashi M.,  
 RA Itzuka H., Miyakawa Y., Mayumi M.;  
 RA "The entire nucleotide sequences of two distinct TT virus (TTV)  
 RA isolates (TJN01 and TJN02) remotely related to the original TTV  
 RA isolates.";  
 RL Arch. Virol. 145:1543-1559(2000).  
 DR EMBL; AB028669; BAA94878.1; -  
 DR InterPro; IPR004219; TTVirus\_Unk.  
 DR Pfam; PF02956; TT\_ORF1.1.  
 SQ SEQUENCE 746 AA; 88561 MW; E0B22953AE764E3E CRC64;

Query Match 28.0%; Score 67.5; DB 12; Length 746;  
 Best Local Similarity 33.3%; Pred. No. 3.2;  
 Matches 11; Conservative 2; Mismatches 5; Indels 15; Gaps 1;

QY 5 WPWMPRRKHAEPEAPIMILKKWMPWRRK 37  
 1 11 111 :||| :|||  
 Db 3 WGMWRRRR-----RMPARRRRR 20

## RESULT 3

09DUC4 ID 09DUC4 PRELIMINARY; PRT; 723 AA.  
 AC 09DUC4;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE ORF1.  
 OS TT virus.  
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.  
 OX NCBI\_TaxID=68887;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MF-TTV9;  
 RA Okamoto H.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MF-TTV9;  
 RA MEDLINE=20534983; PubMed=11080484;  
 RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,  
 RA Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;  
 RA "Species-specific TT viruses in humans and nonhuman primates and their  
 RA phylogenetic relatedness.";  
 RL Virology 277:368-378(2000).  
 DR EMBL; AB041959; BAB19313.1; -  
 DR InterPro; IPR001563; Serine\_carboxypept.  
 DR Pfam; PF02956; TT\_ORF1.1.  
 DR PROSITE; PS00131; CARBOXYPEPT\_SFR\_SER; UNKNOWN.1.  
 SQ SEQUENCE 723 AA; 85393 MW; 232D003098766344 CRC64;

Query Match 27.8%; Score 67; DB 12; Length 723;

Best Local Similarity 34.4%; Pred. No. 3.6;  
 Matches 11; Conservative 1; Mismatches 2; Indels 18; Gaps 1;

QY 6 PWWMPRRKHAEPEAPIMILKKWMPWRRK 37  
 1 11 111 :||| :|||  
 Db 2 PWWMPRR-----WRWRRR 15

## RESULT 4

09IRD8 ID 09IRD8 PRELIMINARY; PRT; 175 AA.  
 AC 09IRD8;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE ORF3.  
 OS TT virus.  
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.  
 OX NCBI\_TaxID=68887;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=L03;  
 RA Liu Z.H., Luo K.X., Hu J., He H.T.;  
 RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF371370; AAK54733.1; -  
 DR InterPro; IPR004219; TTVirus\_Unk.  
 DR Pfam; PF02956; TT\_ORF1.1.  
 SQ SEQUENCE 175 AA; 22073 MW; 5212D7DA3FD72F81 CRC64;

Query Match 26.6%; Score 64; DB 12; Length 175;  
 Best Local Similarity 33.3%; Pred. No. 2.1;  
 Matches 11; Conservative 2; Mismatches 4; Indels 16; Gaps 2;

QY 5 WPWMPRRKHAEPEAPIMILKKWMPWRRK 37  
 1 11 111 :||| :|||  
 Db 3 WSWW-WRRRR-----WMPRRRR 19

## RESULT 5

09D780 ID 09D780 PRELIMINARY; PRT; 49 AA.  
 AC 09D780;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ORF1 (Fragment).  
 OS TT virus.  
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.  
 OX NCBI\_TaxID=68887;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TYM9;  
 RA MEDLINE=20568739; PubMed=11118348;  
 RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,  
 RA Sai T., Sugai Y.;  
 RA "TT virus mRNAs detected in the bone marrow cells from an infected  
 RA individual.";  
 RL Biochem. Biophys. Res. Commun. 279:700-707(2000).  
 DR EMBL; AB050449; BAB19930.1; -  
 DR NON\_TER 49  
 FT SEQUENCE 49 AA; 7225 MW; 1DA6F8F1AB69AA43 CRC64;

Query Match 26.1%; Score 63; DB 12; Length 49;  
 Best Local Similarity 30.3%; Pred. No. 0.79;  
 Matches 10; Conservative 3; Mismatches 4; Indels 16; Gaps 2;

QY 5 WPWMPRRKHAEPEAPIMILKKWMPWRRK 37  
 1 11 111 :||| :|||  
 Db 3 WTWW-WORRRR-----WMPRRR 19

## RESULT 6

09Z8B7

```

* ID 092887 PRELIMINARY; PRT; 192 AA.
AC 092887;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE C1277 SIMILARITY (CPJ0426 protein).
GN CPN0426 OR CPJ0426 OR CPJ327.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CML029;
RA MEDLINE=99206606; PubMed=10192388;
RA Kaldan S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.,
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RA MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Ueterbach T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RA MEDLINE=20330349; PubMed=10871362;
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CML029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL: AE001625; AAD18570.1; -
DR EMBL: AE002194; AAF38182.1; -
DR EMBL: AP002546; BAA98634.1; -
DR TIGR: CP0327; -
KM Complete proteome.
SQ SEQUENCE 192 AA; 21464 MW; 9C38C329AEDDB76F CRC64;

Query Match 26.1%; Score 63; DB 16; Length 192;
Best Local Similarity 42.3%; Pred. No. 3;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 5 WPMWPMRRKHAEPEAPIMILKKWP 30
DB 138 WPMWPMRRKHAEPEAPIMILKKWP 163

RESULT 7
O92881 PRELIMINARY; PRT; 748 AA.
AC O92881;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ORF1.
OS TT virus.
RC TISSUE=LIVER;
RA MEDLINE=87064324; PubMed=3023845;
RA D'Amposio E., Walzkin S.D., Wilney F.R., Saleme A., Furano A.V.;
RT "Structure of the highly repeated, long interspersed DNA family (LINE
RT or L1rn) of the rat."
RL Mol. Cell. Biol. 6:411-424(1986).
DR EMBL: M13100; AAA6046.1; -
DR InterPro: IPR000566; Lipocin_cyFABP.
DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_1.
KM "TT virus mRNAs detected in the bone marrow cells from an infected

```

```

RT Individual.";
RL Biochem. Biophys. Res. Commun. 279:700-707(2000).
DR EMBL: AB050448; BAB19928.1; -
DR InterPro: IPR004219; TTvirus_Unk.
DR Pfam: PF02956; TT_ORF1; 1.
SQ SEQUENCE 748 AA; 88552 MW; D6SCB20CA5CE26F CRC64;

Query Match 26.1%; Score 63; DB 12; Length 748;
Best Local Similarity 30.3%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 4; Indels 16; Gaps 2;

OY 5 WPMWPMRRKHAEPEAPIMILKKWPMWRRK 37
DB 3 WPMWPMRRKHAEPEAPIMILKKWPMWRRK 19

RESULT 8
O91D04 PRELIMINARY; PRT; 750 AA.
AC O91D04;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ORF1.
OS TT virus.
RC TISSUE=LIVER;
RA MEDLINE=21488921; PubMed=11601907;
RA Okamoto H., Nishizawa T., Takahashi M., Asabe S., Tsuda F.,
RA Yoshikawa A.;
RT "Heterogeneous distribution of TT virus of distinct genotypes in
RT multiple tissues from infected humans."
RL Virology 288:358-368(2001).
DR EMBL: AB060592; BAB69900.1; -
DR InterPro: IPR004219; TTvirus_Unk.
DR Pfam: PF02956; TT_ORF1; 1.
SQ SEQUENCE 750 AA; 89223 MW; 616EC86DC3469091 CRC64;

Query Match 26.1%; Score 63; DB 12; Length 750;
Best Local Similarity 30.3%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 4; Indels 16; Gaps 2;

OY 5 WPMWPMRRKHAEPEAPIMILKKWPMWRRK 37
DB 3 WPMWPMRRKHAEPEAPIMILKKWPMWRRK 19

RESULT 9
O63778 PRELIMINARY; PRT; 367 AA.
AC O63778;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Hypothetical 43.7 kDa protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA MEDLINE=87064324; PubMed=3023845;
RA D'Amposio E., Walzkin S.D., Wilney F.R., Saleme A., Furano A.V.;
RT "Structure of the highly repeated, long interspersed DNA family (LINE
RT or L1rn) of the rat."
RL Mol. Cell. Biol. 6:411-424(1986).
DR EMBL: M13100; AAA6046.1; -
DR InterPro: IPR000566; Lipocin_cyFABP.
DR PROSITE: PS00213; LIPOCALIN; UNKNOWN_1.
KM Hypothetical protein.

```

SQ SEQUENCE 367 AA; 43685 MW; 4654499CC185EC3B CRC64;

Query Match 25.7%; Score 62; DB 11; Length 367;

Best Local Similarity 39.4%; Pred. No. 7.4; Mismatches 13; Conservative 4; Indels 12; Gaps 4; 2;

OY 1 ILKKPMPWPR--RKHEAPEAPIMILK-KW 29  
 DB 25 IFSKWCWFNRATCRMRQIDPSLSPCTKLKSKW 57

#### RESULT 10

08V7E2 PRELIMINARY; PRT; 92 AA.  
 AC 08V7E2;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
 DE ORF1 (Fragment).  
 OS TT virus.  
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.  
 OX NCBI\_TaxID=68887;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2184401; PubMed=11855633;  
 RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,  
 RT Okamoto H.;  
 RT "Analysis of the complete genomes of thirteen TT virus variants  
 RT classifiable into the fourth and fifth genetic groups, isolated from  
 RT viremic infants.";  
 RL Arch. Virol. 147:21-41(2002).  
 DR EMBL; AB064615; BAB79374.1; -;  
 DR InterPro: IPR004219; TTVirus-Unk.  
 DR Pfam: PF02956; TT\_ORF1.1.  
 DR NON\_TER 92  
 SQ SEQUENCE 92 AA; 12429 MW; 188D83D05A7B09A CRC64;

Query Match 25.3%; Score 61; DB 12; Length 92;  
 Best Local Similarity 32.3%; Pred. No. 2.5;  
 Matches 10; Conservative 3; Mismatches 2; Indels 16; Gaps 1;

OY 7 WMPWRRKHEAPEAPIMILKKPMPWPRRK 37  
 DB 3 WMYRRR-----PMRPMRRR 17

#### RESULT 11

08TWG0 PRELIMINARY; PRT; 250 AA.  
 AC 08TWG0;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
 DE NAD-dependent protein deacetylase, STR2 family.  
 GN STR2 OR MK1075.  
 OS Methanopyrus kandleri.  
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
 OC Methanopyrus.  
 OX NCBI\_TaxID=2320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=AV19 / DSM 6324 / JCM 9639;  
 RX MEDLINE=21927647; PubMed=11930014;  
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,  
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
 RA Malakh A.G., Koonin E.V., Kozyavkin S.A.;  
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
 RT and monophyly of archaeal methanogens.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
 DR EMBL; AE010396; AAM02288.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 250 AA; 27799 MW; B051994FE5B24E05 CRC64;

Query Match 25.3%; Score 61; DB 17; Length 250;

Best Local Similarity 50.0%; Pred. No. 6.7; Mismatches 10; Conservative 2; Indels 8; Gaps 0;

OY 3 KKPMPWRRKHEAPEAP 22  
 DB 60 KVMWYIMRRRKIAEAEPNP 79

#### RESULT 12

094C18 PRELIMINARY; PRT; 284 AA.  
 AC 094C18;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
 DE Glycine-rich protein LegRPI.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; eusterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA utappa M., Muchhal U.S., Baldwin J.C., Raghothama K.G.;  
 RT "LegRPI: A new member of glycine-rich proteins from tomato  
 RT (Lycopersicon esculentum)."  
 RL Physiol. Plantarum 0:0-0(2001).  
 DR EMBL; AY026037; AAK08984.1; -;  
 SQ SEQUENCE 284 AA; 23434 MW; E81A84C247CB9ED8 CRC64;

Query Match 25.3%; Score 61; DB 10; Length 284;  
 Best Local Similarity 31.2%; Pred. No. 7.6;  
 Matches 10; Conservative 0; Mismatches 4; Indels 18; Gaps 2;

OY 5 WPMWRRKHEAPEAPIMILKKPMPWPRR 36  
 DB 9 WMMWTR-----WLMW-WER 22

#### RESULT 13

08V7I1 PRELIMINARY; PRT; 734 AA.  
 AC 08V7I1;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
 DE ORF1.  
 OS TT virus.  
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.  
 OX NCBI\_TaxID=68887;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=CT43F;  
 RX MEDLINE=2184401; PubMed=11855633;  
 RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,  
 RA Okamoto H.;  
 RT "Analysis of the complete genomes of thirteen TT virus variants  
 RT classifiable into the fourth and fifth genetic groups, isolated from  
 RT viremic infants.";  
 RL Arch. Virol. 147:21-41(2002).  
 DR EMBL; AB064598; BAB79322.1; -;  
 DR InterPro: IPR004219; TTVirus-Unk.  
 DR Pfam: PF02956; TT\_ORF1.1.  
 SQ SEQUENCE 734 AA; 86978 MW; F60E188BC0104A68 CRC64;

Query Match 25.3%; Score 61; DB 12; Length 734;

Best Local Similarity 32.3%; Pred. No. 19;  
Matches 10; Conservative 3; Mismatches 2; Indels 16; Gaps 1;

Oy 7 WMPWRKHAEPEAPIMILKKMPWRRK 37  
Db 3 WMYRRR-----PWRPWRRR 17

## RESULT 14

O9DUC9 PRELIMINARY; PRT; 735 AA.

AC O9DUC9;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE ORF1.

OS TT virus.  
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.  
OX NCBI\_Taxid=68887;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PT-TTV6;  
RA Okamoto H.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.

RC MEDLINE-20534983; PubMed=11080484;  
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,  
RA Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;  
RT "Species-specific TT viruses in humans and nonhuman primates and their  
RT phylogenetic relatedness";  
RL Virology 277:368-378(2000).  
DR EMBL, AB041957; BABI9308.1;  
DR InterPro: IPR004219; TTVirus\_Unk.  
DR Pfam: PF02956; TT\_ORF1.1.

DR SEQUENCE 735 AA; 86132 MW; 9ED818DBE6FA5D3 CRC64;

Query Match 25.3%; Score 61; DB 12; Length 735;  
Best Local Similarity 29.3%; Pred. No. 19;  
Matches 12; Conservative 3; Mismatches 6; Indels 20; Gaps 3;

Oy 5 WPM-----WMPRRKHAEPEAPIMILKKMPW---PWRRK 37  
Db 3 WPMRRRRRRRRRR-----PWRRRRRRRRTWRRR 31

## RESULT 15

O98414 PRELIMINARY; PRT; 985 AA.

AC O98414;  
DT 01-FEB-1997 (TREMBlrel. 02, Created)  
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.

OS Ovine lentivirus.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11663;

RN [1]  
RP SEQUENCE FROM N.A.

RC MEDLINE-95135990; PubMed=7834396;  
RA Woodward T.M., Carlson J.O., de la Concha-Bermejo A.,  
RA Demartini J.C.;  
RT "Biological and genetic changes in ovine lentivirus strains following  
RT passage in isogenic twin lambs";  
RL J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 8:124-133(1995).  
RN [2]  
RP SEQUENCE FROM N.A.

RC STRAIN=85/34;  
RA Carlson J.O., Demartini J.C., Mwaengo D.M.;  
RT "Envelope glycoprotein nucleotide sequence and genetic

RT characterization of North American ovine lentiviruses.";  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: U64439; AAR08725.1; -;  
DR InterPro: IPR000328; Env\_GP41.

DR Pfam: PF00517; GP41; 1.  
KW Transmembrane.

SO SEQUENCE 985 AA; 113794 MW; 3197258DBDE3597 CRC64;

Query Match 25.3%; Score 61; DB 15; Length 985;  
Best Local Similarity 23.4%; Pred. No. 25;  
Matches 18; Conservative 7; Mismatches 8; Indels 44; Gaps 5;

Oy 1 ILKK-----WP-----WPKRKHAEPEAE----- 21

Db 163 ILKRYKODWPNTYHWPIMQENMKRWKENEREYKGTNTKEDIDDLAKIRGRFC 222

Oy 22 ---PIMLK--KWPWP 33

Db 223 VPPFPALKCTKRCWYP 239

Search completed: December 11, 2002, 15:39:39  
Job time : 30 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:34:48 ; Search time 11 Seconds

(without alignments)  
139.511 Million cell updates/sec

Title: US-09-444-281-35-27-35

Perfect score: 241  
Sequence: 1 ILKKMPMPWRKHEAPEEPIMLKKMPMPWRK 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	32.4	144	1 INDC_BOVIN	P33046 bos taurus
2	62	25.7	314	-1 YMA3_BACST	O45633 bacillus st
3	59.5	24.7	2290	1 P03304 EMCV	P03304 encephalomy
4	57	23.7	711	1 MMLA-STRCO	O53902 streptomyc
5	55.5	23.0	55	1 ATP8_ANAPL	P50655 anas platyr
6	55.5	23.0	942	1 ENV_CAEVG	P31627 caprine art
7	55.5	23.0	990	1 ENV_OMVVS	P16899 ovine lenti
8	55	22.8	715	1 YD55_MYCTU	O11025 mycobacteri
9	54	22.4	144	1 AP22_APIME	P35881 apis mellif
10	54	22.4	469	1 STCL_MYCTU	P96862 mycobacteri
11	54	22.4	473	1 STCL_MYCTU	P57990 mycobacteri
12	54	22.4	1173	1 VGL2_CVH22	P15423 human coron
13	53.5	22.2	1042	1 COR1_HUMAN	O9y9q5 homo sapien
14	53.5	22.2	1113	1 COR1_HUMAN	O9y9q5 mus musculu
15	53	22.0	424	1 FDC6_SOYBN	P46828 glycine max
16	52.5	21.8	55	1 ATP8_AYTAM	O9xkz5 aythya amer
17	52.5	21.8	257	1 E434_ADE40	O64665 human adeno
18	52.5	21.8	691	1 YHOG_ECOLI	P03379 visna lenti
19	52.5	21.8	982	1 ENV_VILV	P35954 visna lenti
20	52.5	21.8	983	1 ENV_VILV	P35954 visna lenti
21	52.5	21.8	991	1 ENV_VILV2	P23423 visna lenti
22	52.5	21.8	1154	1 VGL2_IBVD2	P12722 avian infec
23	52.5	21.8	1162	1 VGL2_IBVD	P11223 avian infec
24	52.5	21.8	1162	1 VGL2_IBVD	P12650 avian infec
25	52.5	21.8	1162	1 VGL2_IBVD	P12651 avian infec
26	52.5	21.8	1163	1 VGL2_IBVD	P05135 avian infec
27	52	21.6	68	1 Y121_BPT4	O02405 bacterioph
28	51.5	21.4	162	1 DSRB_NEIMA	O9jfac6 neisseria m
29	51.5	21.4	162	1 DSRB_NEIMA	O9jfac6 neisseria m
30	51.5	21.4	173	1 NUGC_ANASP	O44241 anabaena sp
31	51.5	21.4	443	1 FDC6_BRANA	P48627 brassica na
32	51.5	21.4	448	1 FDC6_BRANA	P46312 arabidopsis
33	51.5	21.4	989	1 ENV_VILV1	P23422 visna lenti

34	51.5	21.4	1262	1 MYO6_HUMAN	O9un54 homo sapien
35	51.5	21.4	1265	1 MYO6_MOUSE	O64331 mus musculu
36	51	21.2	126	1 YDQ3_MYCTU	O11013 mycobacteri
37	51	21.2	196	1 YAO5_SCHPO	O09677 schizosacch
38	51	21.2	295	1 YDQ3_ECO57	O8x627 escherichia
39	51	21.2	295	1 YDQ3_ECO57	P76213 escherichia
40	51	21.2	556	1 MEND_ECOLI	P17109 e menaquin
41	51	21.2	2292	1 P04G_EMCVB	P17593 encephalomy
42	51	21.2	2292	1 P04G_EMCVD	P17594 encephalomy
43	50.5	21.0	293	1 YDQ3_SALTI	O82695 salmonella
44	50.5	21.0	293	1 YDQ3_SALTY	O82p06 salmonella
45	50.5	21.0	469	1 GATR_THETH	O91cx2 thermus the

## ALIGNMENTS

RESULT 1	ID	INDC_BOVIN	STANDARD:	PRT:	144 AA.
AC	P33046;				
DT	01-OCT-1993 (Rel. 27, Created)				
DT	01-OCT-1993 (Rel. 27, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	Indolicidin precursor.				
OS	Bos taurus (bovine).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Bone marrow;				
RX	MEDLINE=92392368; PubMed=1520337;				
RA	del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;				
RT	"CDNA cloning of the neutrophil bactericidal peptide indolicidin.";				
RL	Biochem. Biophys. Res. Commun. 187:467-472(1992).				
RN	[2]				
RP	SEQUENCE OF 131-143.				
RC	TISSUE=Neutrophils;				
RX	MEDLINE=92165771; PubMed=1537821;				
RA	Selsted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.;				
RT	Cullor J.S.;				
RL	"Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.";				
J.	Biol. Chem. 267:4292-4295(1992).				
CC	-1- FUNCTION: POTENT MICROBICIDAL ACTIVITY, ACTIVE AGAINST STAPHYLOCOCCUS AUREUS AND ESCHERICHIA COLI.				
CC	-1- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.				
CC	-1- PTM: ELASTASE MIGHT BE RESPONSIBLE FOR ITS MATURATION.				
CC	-1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.				
CC	THIS SWISS-PROT ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION - THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL ENTITIES REQUIRES A LICENSE AGREEMENT (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	EMBL: X67340; CAA47755.1; -				
DR	PIR: JCI222; JCI222.				
DR	PIR: A42387; A42387.				
DR	InterPro: IPR001894; Cathelicidin.				
DR	Pfam: PF000666; Cathelicidins; 1.				
DR	ProDom: PD001838; Cathelicidin; 1.				
DR	PROSITE: PS00946; CATHELICIDINS_1; 1.				
DR	PROSITE: PS00947; CATHELICIDINS_2; 1.				
KW	Antibiotic; Amidation; Signal.				
FT	SIGNAL	1	29		POTENTIAL.
FT	PROPEP	30	130		
FT	PEPTIDE	131	143		INDOLICIDIN.
FT	MOD_RES	30	30		PYRROLIDONE CARBOXYLIC ACID (BY

```

FT DISULFID      85       96          BY SIMILARITY.
FT DISULFID     107      124          BY SIMILARITY.
FT MOD_RES      143      143          AMIDATION (G-144 PROVIDE AMIDE GROUP).
SQ SEQUENCE     144 AA; 16479 MW; E3B1CBBE5C09911 CRC64;

Query Match
Best Local Similarity   32.4%; Score 78; DB 1; Length 144;
Matches    10; Conservative    4; Mismatches    4; Indels    0; Gaps    0;

QY      19 EAEPIMILKKPMPWPMWR 36
        | : :: ||| |||||
Db      126 ELQSYVLPKMPMPWPMWR 143

RESULT_2
YMA3_BACST
ID YMA3_BACST STANDARD; PRT; 314 AA.
AC Q45633;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 35.7 kDa protein in maaA 3'-region (ORF3).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 7953;
RX MEDLINE=94247374; PubMed=8190087;
RA Liang E.C., Ferenc T.;
RT "Molecular cloning of a maltose transport gene from Bacillus
RL Mol. Gen. Genet. 243:343-352(1994)."
-- -1- SIMILARITY: BELONGS TO THE UPF0097 FAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between
CC CC the Swiss Institute of Bioinformatics and the EMBL outstation ~
CC CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isdb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
DR DR EMBL; L13418; AAA71981.1; -
DR InterPro; IPR004175; 2_5_ligase.
KW Pfam; PF02834; 2_5_ligase; 2.
KW Hypothetical protein.
SQ SEQUENCE 314 AA; 35735 MW; B54E25FD3F72BFCA CRC64;

Query Match
Best Local Similarity   25.7%; Score 62; DB 1; Length 314;
Matches    17; Conservative    5; Mismatches   13; Indels   14; Gaps    3;

QY      2 LKKWPMPRRKRHEAPEA-----EP----IWLKKWPMP-----WRR 36
        ||::| : : ||: | | | | | | | | | | | | | | | | | | | | |
Db      222 LKQWMEQAKKAWEARDADGARREPGETWDFLLDKPSWERLIGRWRR 270

RESULT_3
POLG_EMCV
ID POLG_EMCV STANDARD; PRT; 2290 AA.
AC P03304;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein (Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; P3A; Genome-linked protein VP6; Picornain 3C DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D DE (EC 2.7.7.48)).
OS Encephalomyocarditis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Cardiovirus.
```

```

XN [1] NCBI_TaxID=12104;
RN RN
RP SEQUENCE FROM N.A.
RX MEDLINE=84169586; PubMed=6324136;
RA Palmenberg A.C., Kirby E.M., Janda M.R., Drake N.L., Duke G.M.,
RA Potratz K.F., Collett M.S.;
RT "The nucleotide and deduced amino acid sequences of the
RL encephalomyocarditis viral polypeptide coding region.";
Nucleic Acids Res. 12:2969-2985(1984).
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -1- CATALYTIC ACTIVITY: Selective cleavage of Glu-1-gly bond in the
CC polyovirin polypeptide. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: "THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4."
CC -1- PMM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb.sib.ch).
CC -----
CC EMBL; X00463; CAA25152.1; -.
DR PIR; A03906; GNMEV.
DR HSSP; P12296; 2MEV.
DR MEROPS; CO3.009; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF000073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; Coat protein; Core protein; Transferase;
RW Polyprotein; Coat protein; RNA_helicase; 1.
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
FT PROPEP 1 67 LEADER PEPTIDE.
FT CHAIN 68 136 COAT PROTEIN VP4 (RHO).
FT CHAIN 137 391 COAT PROTEIN VP2 (BETA).
FT CHAIN 392 622 COAT PROTEIN VP3 (GAMMA).
FT CHAIN 623 910 COAT PROTEIN VP1 (ALPHA).
FT CHAIN 911 1056 CORE PROTEIN P2A (G).
FT CHAIN 1057 1192 CORE PROTEIN P2B (I).
FT CHAIN 1193 1517 CORE PROTEIN P2C (F).
FT CHAIN 1518 1605 CORE PROTEIN P3A.
FT CHAIN 1606 1625 GENOME-LINKED PROTEIN VP6 (H).
FT CHAIN 1626 1830 PICORNAVIRIN 3C (P22).
FT CHAIN 1831 2290 MYRI-DIRECTED RNA POLYMERASE P3D (E).
FT LIPID 68 68 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1784 1784 PROTEINASE (POTENTIAL).
FT ACT_SITE 1802 1802 PROTEASE (POTENTIAL).
SQ SEQUENCE 2290 AA: 255756 MW; 26BC81BB7CF68CB5 CRC64;

Query Match 24.7%; Score 59.5; DB 1; Length 2290;
Best Local Similarity 26.2%; Pred. No. 14;
Matches 11; Conservative 6; Mismatches 8; Indels 17; Gaps 1;

OY 6 PWMWRKRKHEAEPEEAET-----MILKKWP 30
DB 967 PWMWRKNTYQAVLRAPCRVTMDIYYKRVRRPLRYOKEMP 1008

RESULT 4
MMLA_STRCO STANDARD; PRT; 711 AA.
AC MLA_STRCO
DT 30-MAY-2000 (Rel. 39, Created)

```

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DR Putative membrane protein actin-3
GN ACT11-3 OR SC05084 OR SCBAC2861.10.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91347376; PubMed=1878971;
RA Fernandez-Moreno M.A., Caballero J.L., Hopwood D.A., Malpartida F.;
RT "The act cluster contains regulatory and antibiotic export genes,
RT direct targets for translational control by the blaA tRNA gene of
RT Streptomyces."
RL Cell 66:769-780(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=1200953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Watzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE MRP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M64683; AAA26691.1; -.
DR EMBL: AL593842; CAC44197.1; -.
DR InterPro: IPR004707; Act11.
DR InterPro: IPR000731; HMGCR/patch_5TM.
DR Pfam: PF03176; MMP1. 2.
DR TIGRPFAM: TIGR00833; act11. 1.
DR PROSITE: PS50156; SSD. 2.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 14 34
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 235 255 POTENTIAL.
FT TRANSMEM 281 301 POTENTIAL.
FT TRANSMEM 313 333 POTENTIAL.
FT TRANSMEM 369 389 POTENTIAL.
FT TRANSMEM 516 536 POTENTIAL.
FT TRANSMEM 540 560 POTENTIAL.
FT TRANSMEM 573 593 POTENTIAL.
FT TRANSMEM 623 643 POTENTIAL.
FT TRANSMEM 645 665 POTENTIAL.
SQ SEQUENCE 711 AA; 74862 MW; A5466BDBABED1B6 CRC64;

Query Match 23.7%; Score 57; DB 1; Length 711;
Best Local Similarity 47.4%; Pred. No. 9.1;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

RESULT 5
ID ATP8_ANAPL STANDARD; PRT; 55 AA.
AC P50655;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (a6L).
GN ATP8 OR ATP8.
OS Anas platyrhynchos (Domestic duck).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=pekin breed; TISSUE=Liver;
RX MEDLINE=94047124; PubMed=8230253;
RA Ramirez V., Savole P., Morais R.;
RT "Molecular characterization and evolution of a duck mitochondrial
RT genome.";
RL J. Mol. Evol. 37:296-310(1993).
CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF0) SUBUNIT OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) -> ADP + phosphate +
CC H(+) (out).
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L22476; AAA72037.1; -.
DR InterPro: IPR001421; ATPase8_mit.
DR Pfam: PF00895; ATP-synt_8; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 8 24
FT TRANSMEM 55 AA; 6334 MW; 41284BAB1525B99 CRC64;
SQ SEQUENCE 55 AA; 6334 MW; 41284BAB1525B99 CRC64;

Query Match 23.0%; Score 55.5; DB 1; Length 55;
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 9; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

```

RT "Structure and genetic variability of envelope glycoproteins of two  
RT antigenic variants of caprine arthritis-encephalitis lentivirus.";   
RL J. Virol. 65:5744-5750(1991).  
[2]  
RN REVIEWS.  
RA Knowles D.P.;  
RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL, M60855; AAB8709.2; -.  
DR PIR, A41307; VCLJC6.  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KW Glycoprotein; Coat protein; Polypeptide; Transmembrane.  
FT PEPTIDE 1 80  
FT CHAIN 1 80  
FT CHAIN 631 630  
FT DOMAIN 1 630  
FT TRANSMEM 631 659  
FT DOMAIN 660 799  
FT TRANSMEM 800 820  
FT DOMAIN 821 942  
FT CARBOHYD 51 51  
FT CARBOHYD 98 98  
FT CARBOHYD 131 131  
FT CARBOHYD 176 176  
FT CARBOHYD 228 228  
FT CARBOHYD 331 331  
FT CARBOHYD 348 348  
FT CARBOHYD 354 354  
FT CARBOHYD 370 370  
FT CARBOHYD 379 379  
FT CARBOHYD 400 400  
FT CARBOHYD 404 404  
FT CARBOHYD 435 435  
FT CARBOHYD 441 441  
FT CARBOHYD 447 447  
FT CARBOHYD 457 457  
FT CARBOHYD 467 467  
FT CARBOHYD 481 481  
FT CARBOHYD 493 493  
FT CARBOHYD 503 503  
FT CARBOHYD 509 509  
FT CARBOHYD 527 527  
FT CARBOHYD 534 534  
SQ SEQUENCE 942 AA; 108437 MW; 5B0ELF0F3D535F4A CRC64;  
Query Match 23.0%; Score 55.5; DB 1; Length 942;  
Best Local Similarity 21.4%; Pred. No. 18;  
Matches 15; Conservative 7; Mismatches 9; Indels 39; Gaps 4;  
OY 3 KKWPW-----WP-----WRKKHAEPEAF-----PIMI 25  
DB 143 EWWPNNTYHWPLQWMEENVYWLKENIAENKKRKNSTKGIEELLAGTIRGRCPVYPAL 202  
OY 26 LK--KWPMP 33  
DB 203 LKCTKWCWYP 212

DR 01-JUN-1994 (Rel. 29, Last annotation update)  
DE ENV polypeptide precursor (Coat polypeptide).  
GN ENV  
OS Ovine lentivirus (strain SA-OMV).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11664;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90223989; PubMed=2158181;  
RA Querat G., Audoly G., Sonigo P., Vigne R.;  
RT "Nucleotide sequence analysis of SA-OMV, a virus-related ovine  
RT lentivirus: phylogenetic history of lentiviruses.";   
RL Virology 175:434-447(1990).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL, M34193; AAA6783.1; -.  
DR EMBL, M31646; AAA66817.1; -.  
DR PIR, G46335; G46335.  
DR HIV, M34193; ENV5OMVSACG.  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KW Glycoprotein; Coat protein; Polypeptide; Transmembrane.  
FT PEPTIDE 1 101  
FT CHAIN 102 662  
FT CHAIN 663 990  
FT TRANSMEM 842 863  
FT CARBOHYD 141 141  
FT CARBOHYD 162 162  
FT CARBOHYD 207 207  
FT CARBOHYD 259 259  
FT CARBOHYD 299 299  
FT CARBOHYD 363 363  
FT CARBOHYD 386 386  
FT CARBOHYD 402 402  
FT CARBOHYD 413 413  
FT CARBOHYD 434 434  
FT CARBOHYD 438 438  
FT CARBOHYD 469 469  
FT CARBOHYD 474 474  
FT CARBOHYD 480 480  
FT CARBOHYD 490 490  
FT CARBOHYD 500 500  
FT CARBOHYD 514 514  
FT CARBOHYD 526 526  
FT CARBOHYD 536 536  
FT CARBOHYD 542 542  
FT CARBOHYD 550 550  
FT CARBOHYD 560 560  
FT CARBOHYD 567 567  
FT CARBOHYD 703 703  
FT CARBOHYD 771 771  
FT CARBOHYD 778 778  
FT CARBOHYD 794 794  
SQ SEQUENCE 990 AA; 114498 MW; 279B816B55614F3 CRC64;  
Query Match 23.0%; Score 55.5; DB 1; Length 990;  
Best Local Similarity 20.0%; Pred. No. 19;  
Matches 14; Conservative 9; Mismatches 8; Indels 39; Gaps 4;  
OY 3 KKWPW-----WP-----WRKKHAEPEAF-----PIMI 25  
DB 174 QEWPNNTYHWPLQWMEENVYWLKENIAENKKRKNSTKGIEELLAGTIRGRCPVYPAL 233  
OY 26 LK--KWPMP 33  
DB 234 LKCEWMCWYP 243

```

RESULT 8
ID YD55_MYCTU STANDARD; PRT; 715 AA.
AC 011025;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV1355C.
GN RV1355C OR MT1398 OR MTCY02B10.19C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z75555; CAA99988.1; -
DR EMBL: AE007012; AAK45661.1; ALT_INIT.
DR TIGR: MT1398; -
DR Tuberculin; RV1355C; -
DR InterPro: IPR000594; Thif_domain.
DR Pfam: PF00899; Thif; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 715 AA; 78181 MW; 455495248A56041C CRC64;
Query Match 22.8%; Score 55; DB 1; Length 715;
Best Local Similarity 38.1%; Pred. No. 16;
Matches 8; Conservative 7; Mismatches 4; Indels 2; Gaps 1;
OY 18 PEAPEIMI--LKKPMPMPRR 36
ID 53 PDPDPDLLEAKRWAYPMWR 73
Db 53 PDPDPDLLEAKRWAYPMWR 73

```

```

RESULT 9
AP22_APIME STANDARD; PRT; 144 AA.
ID AP22_APIME STANDARD; PRT; 144 AA.
AC P35361; P11525; P11526;
DT 01-OCT-1989 (Rel. 12, Created)

```

```

DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Apidaecin precursor, type 22.
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Apoidea; Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93223697; PubMed=8467807;
RA Casteels-P., Ampe C., Capaci T., Casteels P., Tempst P.;
RT "Apidaecin multipetide precursor structure: a putative mechanism for
RT amplification of the insect antibacterial response."
RL EMBO J. 12:1569-1578(1993).
RN [2]
RP SEQUENCE (APIDACIN IA/TB).
RC TISSUE=Hemolymph;
RX MEDLINE=90005446; PubMed=2676519;
RA Casteels P., Ampe C., Jacobs F., Vaek M., Tempst P.;
RT "Apidaecins: antibacterial peptides from honeybees."
RL EMBO J. 8:2387-2391(1989).
CC -1- FUNCTION: APIDACIN HAVE BACTERICIDAL ACTIVITY; PREDOMINANTLY
CC AGAINST GRAM-NEGATIVE BACTERIA. THEY SEEM TO INTERFERE WITH CELL
CC PROPAGATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X72576; CAA51168.1; -
DR PIR: S05383; S05383.
DR PIR: S06675; S06675.
DR PIR: S35331; S35331.
DR InterPro: IPR004828; Apidaecin.
DR Pfam: PF00807; Apidaecin; 4.
KW Insect immunity; Antibiotic; Hemolymph; Signal; Multigene family;
KW Cleavage on pair of basic residues; Repeat.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 35 42
FT PEPTIDE 43 60 APIDACIN IB.
FT PROPEP 63 70
FT PEPTIDE 71 88 APIDACIN IB.
FT PROPEP 91 98
FT PEPTIDE 99 116 APIDACIN IB.
FT PROPEP 119 126
FT PEPTIDE 127 144
SQ SEQUENCE 144 AA; 16539 MW; 6FALAD74CB77108D CRC64;
Query Match 22.4%; Score 54; DB 1; Length 144;
Best Local Similarity 64.7%; Pred. No. 4;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
OY 6 PMPMPRRKHEAPEAP 22
ID 26 PTRPTRLRREAEPEAP 42
Db 26 PTRPTRLRREAEPEAP 42

```

```

RESULT 10
SYCL_MYCTU STANDARD; PRT; 469 AA.
ID SYCL_MYCTU STANDARD; PRT; 469 AA.
AC P96862;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cysteiny1-trNA synthetase I (EC 6.1.1.16) (Cysteine--trNA ligase 1)
DE (CysRS 1).
GN CYSS1 OR CYSS OR RV3580C OR MT3686 OR MTCY06G11.27C.
OS Mycobacterium tuberculosis.

```



DE E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein).  
GN S.  
OS Human coronavirus (strain 229E).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
OC Coronaviridae; Coronavirus.  
OX NCBI\_TaxId=11137;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90264837; PubMed=2345367;  
RA Raabe T., Scheile-Prinz B., Siddell S.G.;  
RT "Nucleotide sequence of the gene encoding the spike glycoprotein of  
RT human coronavirus HCV 229E.";  
RT J. Gen. Virol. 71:1065-1073(1990).  
CC - FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS  
CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION  
CC AND IN STINGYDUM FORMATION.  
CC - SUBCELLULAR LOCATION: Type I membrane protein.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X16816; CAA34723.1; -  
CC PIR: A34766; VGIHHC.  
CC InterPro: IPR002551; Corona\_S1.  
CC InterPro: IPR002552; Corona\_S2.  
CC Pfam: PF01600; Corona\_S1; 1.  
CC Pfam: PF01601; Corona\_S2; 1.  
KM Glycoprotein; Envelope protein; Transmembrane; Signal.  
FT SIGNAL 1 15  
FT CHAIN 16 1173  
FT DOMAIN 16 1115  
FT TRANSMEM 1116 1135  
FT DOMAIN 1136 1173  
FT DOMAIN 1136 1157  
FT CARBOHYD 23 23  
FT CARBOHYD 62 62  
FT CARBOHYD 98 98  
FT CARBOHYD 147 147  
FT CARBOHYD 171 171  
FT CARBOHYD 176 176  
FT CARBOHYD 220 220  
FT CARBOHYD 243 243  
FT CARBOHYD 326 326  
FT CARBOHYD 333 333  
FT CARBOHYD 440 440  
FT CARBOHYD 464 464  
FT CARBOHYD 518 518  
FT CARBOHYD 538 538  
FT CARBOHYD 542 542  
FT CARBOHYD 568 568  
FT CARBOHYD 581 581  
FT CARBOHYD 587 587  
FT CARBOHYD 663 663  
FT CARBOHYD 671 671  
FT CARBOHYD 930 930  
FT CARBOHYD 1015 1015  
FT CARBOHYD 1020 1020  
FT CARBOHYD 1037 1037  
FT CARBOHYD 1049 1049  
FT CARBOHYD 1061 1061  
FT CARBOHYD 1066 1066  
FT CARBOHYD 1076 1076  
FT CARBOHYD 1082 1082  
FT CARBOHYD 1096 1096  
SQ SEQUENCE 1173 AA; 128639 MW; B9CA9A1A796B3DD CRC64;  
Query Match 22.4%; Score 54; DB 1; Length 1173;  
Best Local Similarity 85.7%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 4 KMFMPW 10  
Db 1113 KMFMPW 1119  
RESULT 13  
ID CORL\_HUMAN STANDARD; PRT; 1042 AA.  
AC Q9Y505; Q9YHYZ;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Atrial natriuretic peptide-converting enzyme (EC 3.4.21.-) (pro-ANP-  
DE converting enzyme) (Corin) (Heart specific serine proteinase ATC2).  
GN CRN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE-Heart;  
RX MEDLINE=99262646; PubMed=10329693;  
RA Yan W., Sheng N., Seto M., Morser J., Wu Q.;  
RT "Corin, a mosaic transmembrane serine protease encoded by a novel cDNA  
RT from human heart.";  
RT J. Biol. Chem. 274:14926-14935(1999).  
RN [2]  
RP SEQUENCE OF 734-1040 FROM N.A.  
RX TISSUE-Heart;  
RX MEDLINE=20534769; PubMed=11082206;  
RA Hooper J.D., Scarnan A.L., Clarke B.E., Normyle J.F., Antalis T.M.;  
RT "Localization of the mosaic transmembrane serine protease corin to  
RT heart myocytes.";  
RT Eur. J. Biochem. 267:6931-6937(2000).  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=20539740; PubMed=10880574;  
RA Yan W., Wu F., Morser J., Wu Q.;  
RT "Corin, a transmembrane cardiac serine protease, acts as a pro-atrial  
RT natriuretic peptide-converting enzyme.";  
RT Proc. Natl. Acad. Sci. U.S.A. 97:8525-8529(2000).  
CC - FUNCTION: CONVERTS PRO-ANP TO ANP. CLEAVES PRO-ANP SPECIFICALLY  
CC BETWEEN ARG-123 AND SER-124.  
CC - SUBCELLULAR LOCATION: Type II membrane protein.  
CC - TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART. EXPRESSED IN HEART  
CC MYOCYTES.  
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC - SIMILARITY: CONTAINS 7 LDL-RECEPTOR CLASS A DOMAINS.  
CC - SIMILARITY: CONTAINS 2 FRIZZLED (FZ) DOMAINS.  
CC - SIMILARITY: CONTAINS 1 SRCR DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF13845; AAD31850.1; -  
CC EMBL: AF13248; AAF21966.1; -  
CC HSSP: P00763; IDPO.  
CC MEROPS: S01.019; -  
CC MIM: 605236; -  
CC InterPro: IPR001314; Chymotrypsin.  
CC InterPro: IPR000024; Fz\_domain.  
CC InterPro: IPR002172; LDL\_recept\_A.  
CC InterPro: IPR001254; Ser\_protease\_Try.  
CC InterPro: IPR001190; Src\_receptor.  
CC Pfam: PF00057; Idl\_recept\_a; 6.





```

FT DOMAIN 373 408 IDL-RECEPTOR CLASS A 2.
FT DOMAIN 409 445 IDL-RECEPTOR CLASS A 3.
FT DOMAIN 446 483 IDL-RECEPTOR CLASS A 4.
FT DOMAIN 518 641 F2.
FT DOMAIN 647 682 IDL-RECEPTOR CLASS A 5.
FT DOMAIN 683 721 IDL-RECEPTOR CLASS A 6.
FT DOMAIN 722 757 IDL-RECEPTOR CLASS A 7.
FT DOMAIN 758 853 SRCR.
FT DOMAIN 869 1113 SERINE PROTEASE.
FT ACT_SITE 910 910 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 959 959 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1052 1052 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 857 979 BY SIMILARITY.
FT DISULFID 895 911 BY SIMILARITY.
FT DISULFID 1022 1037 BY SIMILARITY.
FT DISULFID 1048 1077 BY SIMILARITY.
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 519 519 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 719 719 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 765 765 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 828 828 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 970 970 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1113 AA; 122984 MW; B845B2C5F20DD8EC CRC64;

```

```

Query Match 22.2%; Score 53.5; DB 1; Length 1113;
Best Local Similarity 36.0%; Pred. No. 38;
Matches 9; Conservative 5; Mismatches 8; Indels 3; Gaps 1;

```

OY 8 WPMRRKHEAPEAE--PIMILKKW 29

DB 880 WPMQCSLQSEPSGHCVCVLIARKKW 904

```

RESULT 15
FDCG_SQYBN STANDARD; PRT; 424 AA.
AC P48628;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Omega-6 fatty acid desaturase, chloroplast precursor (EC 1.14.99.-).
OS Glycine max (soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=94345008; PubMed=8066133;
RA Hitz W.D., Carlson T.J., Booth J.R. Jr., Kinney A.J., Stecca K.L.,
RA Yadev N.S.;
RT "Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA
RT and its expression in a cyanobacterium.";
RL Plant Physiol. 105:635-641(1994).
CC -1- FUNCTION: CHLOROPLAST OMEGA-6 FATTY ACID DESATURASE INTRODUCES
CC THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 16:3 AND 18:3 FATTY
CC ACIDS. IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT
CC TO USE FERREDOXIN AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS
CC ESTERIFIED TO GALACTOLIPIDS, SULFOLIPIDS AND PHOSPHATIDYLGLYCEROL.
CC -1- PATHWAY: Polyunsaturated fatty acid biosynthesis.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).

```

```

CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/OR BE INVOLVED IN METAL ION BINDING.
CC -1- SIMILARITY: BELONGS TO THE FATTY ACID DESATURASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: I29215; AA50158.1; -
CC InterPro: IPR001225; FA_desaturase.
CC Pfam: PF00487; FA_desaturase.1.
CC ProDom: PD001081; FA_desaturase.1.
CC Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
CC Transf. 1
CC TRANSIT 1 63 CHLOROPLAST (BY SIMILARITY).
CC CHAIN 64 424 OMEGA-6 FATTY ACID DESATURASE.
CC DOMAIN 165 169 HISTIDINE BOX-1.
CC DOMAIN 201 205 HISTIDINE BOX-2.
CC DOMAIN 361 365 HISTIDINE BOX-3.
SQ SEQUENCE 424 AA; 49641 MW; 4F3DF52D4B1A2009 CRC64;

```

```

Query Match 22.0%; Score 53; DB 1; Length 424;
Best Local Similarity 34.4%; Pred. No. 17;
Matches 11; Conservative 7; Mismatches 12; Indels 2; Gaps 2;

```

OY 5 WPMWPMRRKHEAPEAEPIIMILKKWMPMP-WR 35

DB 192 YPEPMPFKHNRH-HAKTNLNRDTAMHNPVK 222

Search completed: December 11, 2002, 15:39:03  
Job time : 13 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:38:09 ; Search time 19 seconds  
(without alignments)  
187.209 Million cell updates/sec

Title: US-09-444-281-35-27-35  
Perfect score: 241  
Sequence: 1 ILKKWPMWPMWRKHEAPEEPIMILKKWPMWPMWRK 37

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	32.4	144	1 JCI1222	indolicidin precursor
2	63	26.1	192	2 H86543	hypothetical prote
3	63	26.1	192	2 D72081	conserved hypothet
4	62	25.7	314	2 S43916	hypothetical prote
5	60	24.9	381	2 B87470	hypothetical prote
6	59.5	24.7	114	2 T36208	hypothetical prote
7	59.5	24.7	2280	1 GNNYE	genome polypeptide
8	59	24.5	95	2 E86447	protein F5D14.5 [1
9	59	24.5	485	2 S74708	hypothetical prote
10	58	24.1	107	2 T35634	hypothetical prote
11	58	24.1	513	2 S21976	probable RNA-direc
12	57.5	23.9	376	2 T40591	hypothetical prote
13	57	23.7	711	2 C40046	antibiotic transpo
14	56	23.2	452	2 T28094	hypothetical prote
15	55.5	23.0	527	2 S33068	myosin heavy chain
16	55.5	23.0	990	1 G46335	env polypeptide pr
17	55.5	23.0	1940	2 A59287	myosin heavy chain
18	55	22.8	245	2 C69459	transcription regu
19	55	22.8	352	2 S77448	hypothetical prote
20	55	22.8	715	2 B70741	probable moey prote
21	54.5	22.6	134	2 E72532	hypothetical prote
22	54.5	22.6	340	2 PH0217	reverse transcript
23	54.5	22.6	451	2 S30401	hypothetical prote
24	54.5	22.6	547	2 T54635	hypothetical prote
25	54	22.4	144	2 S35331	apidecin 22 precu
26	54	22.4	187	2 AB1647	hypothetical prote
27	54	22.4	469	2 B70607	probable cysteiny
28	54	22.4	473	2 C86949	probable cysteiny
29	54	22.4	1173	1 VGIHHC	E2 glycoprotein pr

30	53.5	22.2	538	2 B84759	hypothetical prote
31	53.5	22.2	113	2 JE0315	low-density lipopr
32	53	22.0	68	2 E82799	hypothetical prote
33	53	22.0	406	2 H69143	coenzyme P420-redu
34	53	22.0	424	2 T07742	omega-6 desaturase
35	53	22.0	480	2 JC7552	Shb-like adapter p
36	53	22.0	691	2 D71430	hypothetical prote
37	53	22.0	949	2 E75352	glycine cleavage s
38	53	22.0	2292	2 S35961	capsid polypeptide
39	52.5	21.8	55	2 T11026	H+-transporting tw
40	52.5	21.8	228	2 AD3326	chloramphenicol O-
41	52.5	21.8	449	2 C84618	hypothetical prote
42	52.5	21.8	471	2 T50016	transcription fact
43	52.5	21.8	490	2 T21365	hypothetical prote
44	52.5	21.8	518	2 T29589	hypothetical prote
45	52.5	21.8	621	2 S37664	peptidomeric polypro

## ALIGNMENTS

RESULT 1  
JCI1222  
indolicidin precursor - bovine  
N:Alternate names: antimicrobial peptide  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 10-Sep-1999 #sequence: revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: JCI1222; A42387; S25664  
R:del Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.  
Biochem. Biophys. Res. Commun. 187, 467-472, 1992  
A:Title: cDNA cloning of the neutrophil bactericidal peptide indolicidin.  
A:Reference number: JCI1222; MUID:92392368; PMID:1520337  
A:Accession: JCI1222  
A:Molecule type: mRNA  
A:Residues: 1-144 <SAL>  
A:Cross-references: EMBL:X67340; NID:9462; PIDN:CAA47755.1; PID:9463  
A:Experimental source: bone marrow  
R:Selsted, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.Q.; Smith, W.; Cullor, J.S.  
J. Biol. Chem. 267, 4292-4295, 1992  
A:Title: Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.  
A:Reference number: A42387; MUID:92165771; PMID:1537821  
A:Accession: A42387  
A:Molecule type: protein  
A:Residues: 131-143 <SEL>  
A:Experimental source: neutrophils  
A>Note: sequence extracted from NCBI backbone (NCBI:83840)  
C:Superfamily: cathelin; cystatin homology  
C:Keywords: amidated carboxyl end  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:32-129/Domain: cystatin homology <CYS>  
F:30-110/Domain: propeptide #status predicted <PRO>  
F:131-143/Product: indolicidin #status experimental <MAT>  
F:143/Modified site: amidated carboxyl end (Arg) (amide in mature form from following  
Query Match 32.4%; Score 78; DB 1; Length 144;  
Best Local Similarity 55.6%; Pred. No. 0.011;  
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
OY 19 EAEPIMLKKWPMWR 36  
Db 126 ELQSVILPMKKWPMWR 143  
RESULT 2  
H86543  
hypothetical protein CPJ0426 (imported) - Chlamydomonas pneumoniae (strain J138)  
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae  
C>Date: 02-Mar-2001 #sequence: revision 02-Mar-2001 #text\_change 02-Mar-2001  
C:Accession: H86543  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.  
A:Reference number: AB6491; MUID:20330349; PMID:10871362



C>Date: 28-Aug-1985 #sequence\_revision 28-Aug-1985 #text\_change 16-Jul-1999  
 C/Accession: A03906; JN0383  
 R:Palmenberg, A.C.; Kirby, E.M.; Janda, M.R.; Drake, N.L.; Duke, G.M.; Potratz, K.F.; Co  
 Nucleic Acids Res. 12, 2969-2985, 1984  
 A>Title: The nucleotide and deduced amino acid sequences of the encephalomyocarditis vir  
 A/Reference number: A03906; M0ID:84169586; PMID:6524136  
 A/Accession: A03906  
 A:Molecule type: genomic RNA  
 A/Residues: 1-2290 <PA>  
 A/Cross-references: GB:X00463; NID:61034; PIDN:CAA25152.1; PID:g61035  
 R:Belkov, N.A.; Chizhikov, V.E.; Blinov, V.M.; Katglinov, V.A.; Mkrtyukov, N.N.; Gutorov,  
 Bioray. Khim. 10, 274-279, 1994  
 A>Title: Nucleotide sequence of the 3'-terminus of encephalomyocarditis virus RNA.  
 A/Reference number: JN0383; M0ID:85022788; PMID:6091680  
 A/Accession: JN0383  
 A:Molecule type: genomic RNA  
 A/Residues: 1337-1386, 'L', 1398-1517, 'A', 1519-1536, 'E', 1538-1556, 'S', 1558-1611, 'T', 1613-1  
 A/Cross-references: GB:M54935  
 A>Note: the authors translated the codon CAU for residue 713 as Thr and AAC for residue  
 C:Superfamily: foot-and-mouth disease virus genome polypeptide  
 C:Keywords: coat protein; core protein; genome-linked protein; hydrolase; nucleotidyltra  
 F:1-67/Domain: leader peptide #status predicted <LDP>  
 F:68-136/Product: coat protein VP4 #status predicted <VP4>  
 F:137-391/Product: coat protein VP2 #status predicted <VP2>  
 F:392-622/Product: coat protein VP3 #status predicted <VP3>  
 F:623-910/Product: coat protein VP1 #status predicted <VP1>  
 F:911-1056/Product: core protein P2-A #status predicted <P2A>  
 F:1057-1192/Product: core protein P2-B #status predicted <P2B>  
 F:1193-1517/Product: core protein P2-C #status predicted <P2C>  
 F:1518-1605/Product: core protein P3-A #status predicted <P3A>  
 F:1606-1625/Product: genome-linked protein VPg #status predicted <VPg>  
 F:1626-1830/Product: proteinase #status predicted <PTS>  
 F:1831-2290/Product: RNA-directed RNA polymerase #status predicted <RDP>

Query Match 24.7%; Score 59.5; DB 1; Length 2290;  
 Best Local Similarity 26.2%; Pred. No. 38;  
 Matches 11; Conservative 6; Mismatches 8; Indels 17; Gaps 1;

QY 6 PMWPMRKHEAPEAPPI-----MLKKWP 30  
 Db 967 PMWPMRKTYQAVLRAPCRVTMDIYKKRVRPRLPLVQKEMP 1008

RESULT 8  
 E86447  
 protein FSD14.5 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C/Accession: E86447  
 R:Rheologis, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Hultzer, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/Reference number: A86141; M0ID:21016719; PMID:11130712  
 A/Accession: E86447  
 A:Molecule type: preliminary  
 A/Status: preliminary  
 A:Molecule type: DNA  
 A/Residues: 1-95 <SRO>  
 A/Cross-references: GB:AE005172; NID:98920603; PIDN:AAF81325.1; GSPDB:GN00141  
 C/Genetics:  
 A:Gene: FSD14.5  
 A:Map position: 1

Query Match 24.5%; Score 59; DB 2; Length 95;  
 Best Local Similarity 21.3%; Pred. No. 1.6;  
 Matches 10; Conservative 5; Mismatches 4; Indels 28; Gaps 3;

QY 5 WP-----WPMRKRHEAPEAPIMLK----KWPMPW 34  
 Db 46 WPVVVVAVGVGGRNMWMM-----PVLITVGGEMWMMW 81

RESULT 9  
 S74708  
 hypothetical protein slr1306 - Synechocystis sp. (strain PCC 6803)  
 C/Species: Synechocystis sp.  
 A/Variety: PCC 6803  
 C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 C/Accession: S74708  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
 O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
 DNA Res. 3, 109-136, 1996  
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
 s.  
 A/Reference number: S74322; M0ID:97061201; PMID:8905231  
 A/Accession: S74708  
 A/Status: preliminary  
 A:Molecule type: DNA  
 A/Residues: 1-485 <KAN>  
 A/Cross-references: EMBL:D90901; GB:AB001339; NID:9151897; PIDN:BA016859.1; PID:d101  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 24.5%; Score 59; DB 2; Length 485;  
 Best Local Similarity 29.3%; Pred. No. 8.6;  
 Matches 12; Conservative 6; Mismatches 5; Indels 18; Gaps 2;

QY 6 PW-----WPMRKRHEAPEAPIM--ILKK 28  
 Db 41 PMDQWGLMALSGGLVIMRMRRRRHAPPEQKQMLTPVELK 81

RESULT 10  
 T35634  
 hypothetical protein SC6G9.38 - Streptomyces coelicolor  
 C/Species: Streptomyces coelicolor  
 C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C/Accession: T35634  
 R:Seeger, K.D.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A  
 submitted to the EMBL Data Library, June 1999  
 A/Reference number: T35634  
 A/Accession: T35634  
 A/Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A/Residues: 1-107 <SEE>  
 A/Cross-references: EMBL:AL079356; PIDN:CA845629.1; GSPDB:GN00070; SCOEDB:SC6G9.38  
 C/Genetics:  
 A:Experimental source: strain A3(2)  
 A:Gene: SCOEDB:SC6G9.38

Query Match 24.1%; Score 58; DB 2; Length 107;  
 Best Local Similarity 50.0%; Pred. No. 2.3;  
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 KKWPMPMRKRHEAPEAP 22  
 Db 79 KKWPMSKPERPSAEPDAP 98

RESULT 11  
 S21976  
 probable RNA-directed DNA polymerase (EC 2.7.7.49) (clone MH2C) - rat retrotransposon  
 N/Alternate names: reverse transcriptase  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 19-Mar-1997 #sequence\_revision 17-Oct-1997 #text\_change 26-Aug-1999  
 C/Accession: S21976  
 R:Kahre, O.; Ilves, H.; Speck, M.  
 submitted to the EMBL Data Library, August 1991  
 A/Reference number: S16783  
 A/Accession: S21976  
 A:Molecule type: mRNA

A:Residues: 1-513 <KAH>  
 A:Cross-references: EMBL:X61295; NID:956521; PIDN:CAA43593.1; PID:956522  
 A:Experimental source: clone M12C  
 C:Genetics:  
 A:Mobile element: retrotransposon L1  
 C:Superfamily: pol polyprotein  
 C:Keywords: nucleotidyltransferase; polypeptide; reverse transcriptase

Query Match 24.1%; Score 58; DB 2; Length 513;  
 Best Local Similarity 39.4%; Pred. No. 12;  
 Matches 13; Conservative 4; Mismatches 12; Indels 4; Gaps 2;

QY 1 ILKKWPMWPMWPMWPMWPMWPMW 29  
 DB 373 IFSKMCMWPMWPMWPMWPMWPMW 405

## RESULT 12

T40591  
 hypothetical protein SPBC646.15c - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 04-Mar-2000  
 C:Accession: T40591  
 R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, January 1999  
 A:Reference number: Z21938  
 A:Accession: T40591  
 A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-376 <SEE>  
 A:Cross-references: EMBL:AL035216; PIDN:CAA22819.1; GSPDB:GN00067; SPDB:SPBC646.15c  
 A:Experimental source: strain 972h-; cosmid c646  
 C:Genetics:  
 A:Gene: SPDB:SPBC646.15c  
 A:Map position: 2 126/2: 312/2: 350/1  
 A:Introns: 48/1; 126/2: 312/2: 350/1  
 C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC646.15c

Query Match 23.9%; Score 57.5; DB 2; Length 376;  
 Best Local Similarity 37.0%; Pred. No. 10;  
 Matches 10; Conservative 3; Mismatches 7; Indels 7; Gaps 1;

QY 8 WPMWRKHEAPEAPIMILKKWPMWPMW 34  
 DB 236 WPMWRKHEAPEAPIMILKKWPMWPMW 255

## RESULT 13

CA0046  
 antibiotic transport-associated protein actr-3 - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 22-Oct-1999  
 C:Accession: CA0046  
 R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Malpartida, F.  
 Cell 66, 769-780, 1991  
 A:Title: The act cluster contains regulatory and antibiotic export genes, direct targets  
 A:Reference number: A40046; MUID:91347376; PMID:1878971  
 A:Accession: CA0046  
 A:Molecule type: DNA  
 A:Residues: 1-711 <PER>  
 A:Cross-references: GB:M64683; NID:9153143; PIDN:AAA26691.1; PID:9153146

Query Match 23.7%; Score 57; DB 2; Length 711;  
 Best Local Similarity 47.4%; Pred. No. 23;  
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 ILKKWPMWPMWPMWPMWPMWPMW 19  
 DB 332 IFGRWPMWPMWPMWPMWPMWPMW 350

RESULT 14  
 T28094

hypothetical protein ZK899.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
 C:Accession: T28094  
 R:Kershaw, J.  
 submitted to the EMBL Data Library, September 1994  
 A:Reference number: Z20468  
 A:Accession: T28094  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-452 <NTL>  
 A:Cross-references: EMBL:Z37140; PIDN:CAA85502.1; GSPDB:GN00028; CESP:ZK899.2  
 A:Experimental source: clone ZK899  
 C:Genetics:  
 A:Gene: CESP:ZK899.2  
 A:Map position: X  
 A:Introns: 34/3; 143/2; 227/2; 262/3; 380/3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein ZK899.2

Query Match 23.2%; Score 56; DB 2; Length 452;  
 Best Local Similarity 32.3%; Pred. No. 19;  
 Matches 10; Conservative 3; Mismatches 10; Indels 8; Gaps 2;

QY 4 KPMWPMWRKHEAPEAPIMILKKWPMWPMW 34  
 DB 172 KLVWMTWPMWPMWPMWPMWPMWPMW 194

## RESULT 15

S33068  
 myosin heavy chain - fluke (Schistosoma mansoni) (fragment)  
 N:Alternate names: surface antigen, 200K  
 C:Species: Schistosoma mansoni  
 C>Date: 22-Nov-1993 #sequence\_revision 06-Sep-1996 #text\_change 13-Feb-1998  
 C:Accession: S33068  
 R:Soisson, L.M.A.; Masterson, C.P.; Tom, T.D.; McNally, M.T.; Lowell, G.H.; Strand, M.  
 J. Immunol. 149, 3612-3620, 1992  
 A:Title: Induction of protective immunity in mice using a 62-kDa recombinant fragment  
 A:Reference number: A46514; MUID:93056536; PMID:1431131  
 A:Accession: S33068  
 A:Molecule type: mRNA  
 A:Residues: 1-527 <SOI>  
 A:Cross-references: EMBL:X65591  
 A:Note: the authors translated the codon CAA for residue 346 as Lys  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: ATP; surface antigen

Query Match 23.0%; Score 55.5; DB 2; Length 527;  
 Best Local Similarity 30.8%; Pred. No. 25;  
 Matches 12; Conservative 7; Mismatches 9; Indels 11; Gaps 2;

QY 1 ILKKWPMWPMWPMWPMWPMWPMWPMW 28  
 DB 106 VLKMPWPMWPMWPMWPMWPMWPMWPMW 144

Search completed: December 11, 2002, 15:40:25  
 Job time: 21 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:39:44 ; Search time 11 Seconds

(without alignments)  
54.633 Million cell updates/sec

Title: US-09-444-281-35-27-35

Perfect score: 241

Sequence: 1 ILKKWPMWPMRRKHEAPEAPIMILKKWPMWPMRRK 37

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications-AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141.5	58.7	28	10	US-09-030-619-104
2	137.5	57.1	28	10	US-09-030-619-50
3	99	41.1	21	10	US-09-030-619-69
4	98	40.7	21	10	US-09-030-619-52
5	93.5	38.8	20	10	US-09-030-619-51
6	91	37.8	13	10	US-09-030-619-95
7	91	37.8	13	10	US-09-030-619-99
8	91	37.8	14	10	US-09-030-619-72
9	91	37.8	14	10	US-09-030-619-107
10	87.5	36.3	21	10	US-09-030-619-47
11	87	36.1	12	10	US-09-030-619-67
12	86	35.7	12	10	US-09-030-619-112
13	86	35.7	13	10	US-09-030-619-53
14	86	35.7	13	10	US-09-030-619-107
15	86	35.7	13	10	US-09-030-619-109
16	86	35.7	14	10	US-09-030-619-54
17	86	35.7	14	10	US-09-030-619-110
18	85	35.3	12	10	US-09-030-619-73
19	85	35.3	13	10	US-09-030-619-41

	20	85	35.3	13	10	US-09-030-619-103	Sequence 103, App
	21	83.5	34.6	15	10	US-09-030-619-39	Sequence 39, App
	22	83	34.4	12	10	US-09-030-619-44	Sequence 44, App
	23	83	34.4	13	10	US-09-030-619-105	Sequence 105, App
	24	83	34.4	13	10	US-09-030-619-106	Sequence 106, App
	25	83	34.4	20	10	US-09-030-619-24	Sequence 24, App
	26	82.5	34.2	21	10	US-09-030-619-48	Sequence 48, App
	27	82	34.0	13	10	US-09-030-619-58	Sequence 58, App
	28	82	34.0	13	10	US-09-030-619-59	Sequence 59, App
	29	82	34.0	13	10	US-09-030-619-94	Sequence 94, App
	30	82	34.0	13	10	US-09-030-619-102	Sequence 102, App
	31	81	33.6	11	10	US-09-030-619-79	Sequence 79, App
	32	81	33.6	11	10	US-09-030-619-114	Sequence 114, App
	33	78	32.4	11	10	US-09-030-619-78	Sequence 78, App
	34	78	32.4	11	10	US-09-030-619-113	Sequence 113, App
	35	78	32.4	12	10	US-09-030-619-43	Sequence 43, App
	36	77.5	32.2	12	10	US-09-030-619-40	Sequence 40, App
	37	77	32.0	12	10	US-09-030-619-82	Sequence 82, App
	38	76	31.5	12	10	US-09-030-619-82	Sequence 82, App
	39	75	31.1	12	10	US-09-030-619-23	Sequence 23, App
	40	75	31.1	12	10	US-09-030-619-30	Sequence 30, App
	41	75	31.1	12	10	US-09-030-619-111	Sequence 111, App
	42	73.5	30.5	12	10	US-09-030-619-75	Sequence 75, App
	43	73.5	30.5	12	10	US-09-030-619-76	Sequence 76, App
	44	73	30.3	9	10	US-09-030-619-80	Sequence 80, App
	45	73	30.3	10	10	US-09-030-619-81	Sequence 81, App

#### ALIGNMENTS

RESULT 1  
US-09-030-619-104  
; Sequence 104, Application US/09030619B  
; Patent No. US20020035061A1  
; GENERAL INFORMATION:  
; APPLICANT: Krieger, Timothy J.  
; APPLICANT: Taylor, Robert  
; APPLICANT: Ertle, Douglas  
; APPLICANT: Fraser, Janet R.  
; APPLICANT: West, Michael H.P.  
; APPLICANT: Montcol, Patricia J.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TITLE OF INVENTION: INJECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION  
; TITLE OF INVENTION: WITH ANTIBIOTICS  
; FILE REFERENCE: 660081.406  
; CURRENT APPLICATION NUMBER: US/09/030,619B  
; CURRENT FILING DATE: 1998-02-25  
; NUMBER OF SEQ ID NOS: 232  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 104  
; LENGTH: 28  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Cationic Peptide Analogue  
US-09-030-619-104

Query Match 58.7%; Score 141.5; DB 10; Length 28;  
Best Local Similarity 58.8%; Pred. No. 4.8e-11;  
Matches 20; Conservative 2; Mismatches 1; Indels 11; Gaps 1;

QY 4 KRPWPMWPMRRKHEAPEAPIMILKKWPMWPMRRK 37  
:||||||| :|||||||  
Db 3 RWPWPMWPMRRK-----ILRRWPMWPMRRK 25

RESULT 2  
US-09-030-619-50  
; Sequence 50, Application US/09030619B  
; Patent No. US20020035061A1  
; GENERAL INFORMATION:  
; APPLICANT: Krieger, Timothy J.

```

; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNICOL, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indollicidin Analogue
US-09-030-619-50
```

```

Query Match          57.1%; Score 137.5; DB 10; Length 28;
Best Local Similarity 61.8%; Pred. No. 1.4e-10;
Matches 21; Conservative 2; Mismatches 0; Indels 11; Gaps 2;
```

```
QY 4 ILKKPMPWRRKHEAPEAPIMILKKMPWRRK 37
DB 3 RWPMPWRRK-----MTL-RWPMPWRRK 25
```

```

RESULT 3
US-09-030-619-69
; Sequence 69, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNICOL, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indollicidin Analogue
US-09-030-619-69
```

```

Query Match          41.1%; Score 99; DB 10; Length 21;
Best Local Similarity 64.3%; Pred. No. 3.6e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
```

```
QY 1 ILKKPMPWRRKHEAPEAPIMILKK 28
DB 1 ILKKPMPWRRK-----MILKK 18
```

```

RESULT 4
US-09-030-619-52
; Sequence 52, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
```

```

; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNICOL, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indollicidin Analogue
US-09-030-619-52
```

```

Query Match          40.7%; Score 98; DB 10; Length 21;
Best Local Similarity 64.3%; Pred. No. 4.7e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
```

```
QY 1 ILKKPMPWRRKHEAPEAPIMILKK 28
DB 1 ILKKPMPWRR-----MILKK 18
```

```

RESULT 5
US-09-030-619-51
; Sequence 51, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNICOL, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indollicidin Analogue
US-09-030-619-51
```

```

Query Match          38.8%; Score 93.5; DB 10; Length 20;
Best Local Similarity 60.7%; Pred. No. 1.5e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 11; Gaps 1;
```

```
QY 1 ILKKPMPWRRKHEAPEAPIMILKK 28
DB 1 ILKKPMPWRR-----MILKK 17
```

```

RESULT 6
US-09-030-619-95
; Sequence 95, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
```



```
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 13
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-95
```

```
Query Match
Best Local Similarity 37.8%; Score 91; DB 10; Length 13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ILKKPMPMPRRK 13
Db 1 ILKKPMPMPRRK 13
```

```
RESULT 7
US-09-030-619-99
; Sequence 99, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 99
; LENGTH: 13
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-99
```

```
Query Match
Best Local Similarity 37.8%; Score 91; DB 10; Length 13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ILKKPMPMPRRK 13
Db 1 ILKKPMPMPRRK 13
```

```
RESULT 8
US-09-030-619-72
; Sequence 72, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
```

```
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 14
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-72
```

```
Query Match
Best Local Similarity 37.8%; Score 91; DB 10; Length 14;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ILKKPMPMPRRK 13
Db 1 ILKKPMPMPRRK 13
```

```
RESULT 9
US-09-030-619-108
; Sequence 108, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 108
; LENGTH: 14
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-108
```

```
Query Match
Best Local Similarity 37.8%; Score 91; DB 10; Length 14;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ILKKPMPMPRRK 13
Db 1 ILKKPMPMPRRK 13
```

```
RESULT 10
US-09-030-619-47
; Sequence 47, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
```

```

; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 47
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indoliciidin Analogue
US-09-030-619-47

Query Match
Best Local Similarity 36.3%; Score 87.5; DB 10; Length 21;
Matches 15; Conservative 1; Mismatches 0; Indels 9; Gaps 1;

QY 4 LKKWPMWPMRRKHEAPEEPIMLKK 28
Db 3 RWPMPWPMRRK-----IMILKK 18

RESULT 11
US-09-030-619-67
; Sequence 67, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indoliciidin Analogue
US-09-030-619-67

Query Match
Best Local Similarity 36.1%; Score 87; DB 10; Length 12;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKKWPMWPMRRK 13
Db 1 LKKWPMWPMRRK 12

RESULT 12
US-09-030-619-112
; Sequence 112, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 112
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-112

Query Match
Best Local Similarity 35.7%; Score 86; DB 10; Length 12;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPMWPMRR 12
Db 1 ILKKWPMWPMRR 12

RESULT 13
US-09-030-619-53
; Sequence 53, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030.619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indoliciidin Analogue
US-09-030-619-53

Query Match
Best Local Similarity 35.7%; Score 86; DB 10; Length 13;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKWPMWPMRR 12
Db 1 ILKKWPMWPMRR 12

RESULT 14
US-09-030-619-107
; Sequence 107, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
```

; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION  
; FILE OF INVENTION: WITH ANTIBIOTICS  
; FILE REFERENCE: 660081.406  
; CURRENT APPLICATION NUMBER: US/09/030,619B  
; CURRENT FILING DATE: 1998-02-25  
; NUMBER OF SEQ ID NOS: 232  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 107  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Cationic Peptide Analogue  
US-09-030-619-107

Query Match 35.7%; Score 86; DB 10; Length 13;  
Best Local Similarity 41.9%; Pred. No. 7.3e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

OY 7 WMPWRRKHEAPEEPIMILKKMPWMPWRRK 37  
|| |||||  
Db 1 WW-----KKMPWMPWRRK 13

RESULT 15  
US-09-030-619-109  
; Sequence 109, Application US/09030619B  
; Patent No. US20020035061A1  
; GENERAL INFORMATION:  
; APPLICANT: Krieger, Timothy J.  
; APPLICANT: Taylor, Robert  
; APPLICANT: Erfle, Douglas  
; APPLICANT: Fraser, Janet R.  
; APPLICANT: West, Michael H.P.  
; APPLICANT: McNICOL, Patricia J.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION  
; TITLE OF INVENTION: WITH ANTIBIOTICS  
; FILE REFERENCE: 660081.406  
; CURRENT APPLICATION NUMBER: US/09/030,619B  
; CURRENT FILING DATE: 1998-02-25  
; NUMBER OF SEQ ID NOS: 232  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 109  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Cationic Peptide Analogue  
US-09-030-619-109

Query Match 35.7%; Score 86; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 7.3e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ILKKMPWMPWRR 12  
|||  
Db 1 ILKKMPWMPWRR 12

Search completed: December 11, 2002, 15:43:29  
Job time : 11 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:37:49 : Search time 15 seconds

(without alignments)  
72.577 Million cell updates/sec

Title: US-09-444-281-35-27-35

Sequence: 1 ILKKMPWMPWRRKHEAPEEPIMLKKMPWMPWRRK 37

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/prodata/1/1aa/6A\_COMB.pep: \*  
4: /cgn2\_6/prodata/1/1aa/6B\_COMB.pep: \*  
5: /cgn2\_6/prodata/1/1aa/PCPUS\_COMB.pep: \*  
6: /cgn2\_6/prodata/1/1aa/backfilest1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	133.5	55.4	63	4	US-09-099-631A-12
2	99.5	41.3	16	4	US-08-702-054B-38
3	99	41.1	21	4	US-08-915-314-54
4	98	40.7	21	4	US-08-915-314-56
5	93.5	38.8	20	4	US-08-915-314-55
6	91	37.8	13	4	US-08-915-314-30
7	91	37.8	13	4	US-08-915-314-62
8	91	37.8	13	4	US-08-915-314-63
9	91	37.8	13	4	US-08-915-314-64
10	91	37.8	13	4	US-09-042-071-36
11	91	37.8	14	4	US-08-915-314-57
12	89	36.9	15	4	US-08-702-054B-40
13	87.5	36.3	21	4	US-08-915-314-46
14	87	36.1	12	4	US-08-915-314-52
15	86	35.7	12	4	US-08-915-314-74
16	86	35.7	12	4	US-08-702-054B-5
17	86	35.7	13	4	US-08-915-314-51
18	86	35.7	13	4	US-08-915-314-58
19	86	35.7	13	4	US-08-702-054B-34
20	86	35.7	14	4	US-08-915-314-59
21	85.5	35.5	16	4	US-08-702-054B-11
22	85	35.3	12	4	US-08-915-314-69
23	85	35.3	13	4	US-08-915-314-38
24	85	35.3	13	4	US-08-915-314-45
25	85	35.3	13	4	US-08-702-054B-1
26	85	35.3	13	4	US-08-702-054B-17
27	85	35.3	13	4	US-08-702-054B-32

28	83.5	34.6	15	4	US-08-702-054B-39	Sequence 39, Appl
29	83	34.4	12	4	US-08-915-314-24	Sequence 24, Appl
30	83	34.4	13	4	US-08-915-314-49	Sequence 49, Appl
31	83	34.4	13	4	US-08-915-314-50	Sequence 50, Appl
32	83	34.4	13	4	US-08-702-054B-30	Sequence 30, Appl
33	83	34.4	13	4	US-08-702-054B-31	Sequence 31, Appl
34	83	34.4	13	4	US-08-702-054B-35	Sequence 35, Appl
35	83	34.4	20	4	US-08-915-314-47	Sequence 47, Appl
36	82.5	34.2	21	4	US-08-915-314-48	Sequence 48, Appl
37	82	34.0	13	4	US-08-915-314-25	Sequence 25, Appl
38	82	34.0	13	4	US-08-915-314-66	Sequence 66, Appl
39	82	34.0	13	4	US-08-915-314-67	Sequence 67, Appl
40	82	34.0	13	4	US-08-702-054B-33	Sequence 33, Appl
41	81	33.6	11	4	US-08-915-314-75	Sequence 75, Appl
42	80	33.2	14	4	US-08-702-054B-18	Sequence 18, Appl
43	80	33.2	15	4	US-08-702-054B-41	Sequence 41, Appl
44	80	33.2	16	4	US-08-702-054B-2	Sequence 2, Appl
45	79	32.8	17	4	US-08-702-054B-42	Sequence 42, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-099-631A-12
; Sequence 12, Application US/09099631A
; Patent No. 6444645
; GENERAL INFORMATION:
; APPLICANT: Selsled, Michael E.
; TITLE OF INVENTION: Crosslink-stabilized Indollicidin Analogs
; FILE REFERENCE: P-UC 3050
; CURRENT APPLICATION NUMBER: US/09/099,631A
; CURRENT FILING DATE: 1998-06-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-099-631A-12

Query Match      55.4%; Score 133.5; DB 4; Length 63;
Best Local Similarity 64.7%; Pred. No. 2.9e-10;
Matches 22; Conservative 0; Mismatches 5; Indels 7; Gaps 2;

QY 4 KMPWMPWRRKHEAPEEPIMLKKMPWMPWRR 36
Db 11 KMPWMPWRRM-----ARIAMILPKMPWMPWRR 38

RESULT 2
US-08-702-054B-38
; Sequence 38, Application US/08702054B
; Patent No. 6191254
; GENERAL INFORMATION:
; APPLICANT: Falls, Timothy J.
; APPLICANT: Hancock, Robert E. W.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/702,054B  
FILING DATE: 23-AUG-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/002,687  
FILING DATE: 23-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07420/013001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-702-054B-38

Query Match 41.3%; Score 99.5; DB 4; Length 16;  
Best Local Similarity 43.2%; Pred. No. 1.2e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 1 ILKKPWPWRRKHEAPEAPIMILKKPWPWRRK 37  
DB 1 ILKKPWPWRRKHEAPEAPIMILKKPWPWRRK 16

## RESULT 3

US-08-915-314-54  
Sequence 54, Application US/08915314  
Patent No. 6180604

GENERAL INFORMATION:  
APPLICANT: Fraser, Janet R.  
APPLICANT: West, Michael H.P.  
APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INOLICIDIN  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,314  
FILING DATE: 20-AUG-1997

## CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
NAME: No. 6180604tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 660081.405  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: linear  
US-08-915-314-54

Query Match 41.1%; Score 99; DB 4; Length 21;  
Best Local Similarity 64.3%; Pred. No. 1.9e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILKKPWPWRRKHEAPEAPIMILKK 28  
DB 1 ILKKPWPWRRKHEAPEAPIMILKK 18

## RESULT 4

US-08-915-314-56  
Sequence 56, Application US/08915314  
Patent No. 6180604

GENERAL INFORMATION:  
APPLICANT: Fraser, Janet R.  
APPLICANT: West, Michael H.P.  
APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INOLICIDIN  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,314  
FILING DATE: 20-AUG-1997

## CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
NAME: No. 6180604tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 660081.405  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 56:

SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-915-314-56

Query Match 40.7%; Score 98; DB 4; Length 21;  
Best Local Similarity 64.3%; Pred. No. 2.3e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 ILKKPWPWRRKHEAPEAPIMILKK 28  
DB 1 ILKKPWPWRRKHEAPEAPIMILKK 18

## RESULT 5

US-08-915-314-55  
Sequence 55, Application US/08915314  
Patent No. 6180604

GENERAL INFORMATION:  
APPLICANT: Fraser, Janet R.  
APPLICANT: West, Michael H.P.

APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
APPLICANT: Erfile, Douglas  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,314  
FILING DATE: 20-AUG-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: NO. 6180604tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 660081.405  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
US-08-915-314-55  
Query Match 38.8%; Score 93.5; DB 4; Length 20;  
Best Local Similarity 60.7%; Pred. No. 8.8e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 11; Gaps 1;  
OY 1 ILKKPMPWRRKHEAPEAEPIMILKK 28  
Db 1 ILKKPMPWRR-----MILKK 17  
RESULT 6  
US-08-915-314-30  
Sequence 30, Application US/08915314  
Patent No. 6180604  
GENERAL INFORMATION:  
APPLICANT: Fraser, Janet R.  
APPLICANT: West, Michael H.P.  
APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
APPLICANT: Erfile, Douglas  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: NO. 6180604tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 660081.405  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
US-08-915-314-30  
Query Match 37.8%; Score 91; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ILKKPMPWRRK 13  
Db 1 ILKKPMPWRRK 13  
RESULT 7  
US-08-915-314-62  
Sequence 62, Application US/08915314  
Patent No. 6180604  
GENERAL INFORMATION:  
APPLICANT: Fraser, Janet R.  
APPLICANT: West, Michael H.P.  
APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
APPLICANT: Erfile, Douglas  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,314  
FILING DATE: 20-AUG-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: NO. 6180604tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 660081.405  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
FEATURE: Modified-site  
NAME/KEY: 1  
LOCATION: 1  
OTHER INFORMATION: /note="D-Form of Isoleucine"  
US-08-915-314-62

Query Match 37.8%; Score 91; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPWRRK 13  
DB 1 ILKKPMPWRRK 13

RESULT 8  
US-08-915-314-63  
; Sequence 63, Application US/08915314  
; Patent No. 6180604  
; GENERAL INFORMATION:  
; APPLICANT: Fraser, Janet R.  
; APPLICANT: West, Michael H.P.  
; APPLICANT: Krieger, Timothy J.  
; APPLICANT: Taylor, Robert  
; APPLICANT: Effie, Douglas  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN  
; NUMBER OF SEQUENCES: 90  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/915,314  
; FILING DATE: 20-AUG-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6180604tenburg Ph.D., Carol  
; REGISTRATION NUMBER: 39,317  
; REFERENCE/DOCKET NUMBER: 660081.405  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 63:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 13  
; OTHER INFORMATION: /note="D-Form of Lysine"  
US-08-915-314-63

Query Match 37.8%; Score 91; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPWRRK 13  
DB 1 ILKKPMPWRRK 13

RESULT 9  
US-08-915-314-64  
; Sequence 64, Application US/08915314  
; Patent No. 6180604  
; GENERAL INFORMATION:  
; APPLICANT: Fraser, Janet R.

APPLICANT: West, Michael H.P.  
APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
APPLICANT: Effie, Douglas  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESS: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,314  
FILING DATE: 20-AUG-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6180604tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 660081.405  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /note="D-Form of Isoleucine"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 13  
OTHER INFORMATION: /note="D-Form of Lysine"  
US-08-915-314-64

Query Match 37.8%; Score 91; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPWRRK 13  
DB 1 ILKKPMPWRRK 13

RESULT 10  
US-09-042-071-36  
; Sequence 36, Application US/09042071  
; Patent No. 6294372  
; GENERAL INFORMATION:  
; APPLICANT: Burian, Jan  
; APPLICANT: Kay, William W.  
; TITLE OF INVENTION: REPLICATION GENES AND GENE PRODUCTS FROM  
; TITLE OF INVENTION: SMALL CRYPTIC PLASMIDS AND METHODS FOR CONSTRUCTING  
; TITLE OF INVENTION: CONTROLLED-REPLICATION PLASMID VECTORS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104



```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/042,071
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 660081.407
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-042-071-36
;
Query Match
37.8%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPWRRK 13
Db 1 ILKKPMPWRRK 13

RESULT 11
US-08-915-314-57
; Sequence 57, Application US/08915314
; Patent No. 6180604
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,314
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6180604tenburg, Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
```

```

;
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-915-314-57
;
Query Match
37.8%; Score 91; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPWRRK 13
Db 1 ILKKPMPWRRK 13

RESULT 12
US-08-702-054B-40
; Sequence 40, Application US/08702054B
; Patent No. 6191254
; GENERAL INFORMATION:
; APPLICANT: Falls, Timothy J.
; APPLICANT: Hancock, Robert E. W.
; APPLICANT: Gough, Monisha
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
; TITLE OF INVENTION: AND METHODS OF SCREENING FOR THE SAME
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,054B
; FILING DATE: 23-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/002,687
; FILING DATE: 23-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07420/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-702-054B-40
;
Query Match
36.9%; Score 89; DB 4; Length 15;
Best Local Similarity 43.8%; Pred. No. 2.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

QY 1 ILKKPMPWRRKHAPEAPEPIILKKPWW 32
Db 1 ILKKPMPWRRK-----HW 14

RESULT 13
US-08-915-314-46
; Sequence 46, Application US/08915314
; Patent No. 6180604
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
```

APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
APPLICANT: Erfile, Douglas  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESS: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,314  
FILING DATE: 20-AUG-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6180604tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 660081.405  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
FAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-915-314-46  
Query Match 36.3%; Score 87.5; DB 4; Length 21;  
Best Local Similarity 60.0%; Pred. No. 5.2e-05;  
Matches 15; Conservative 1; Mismatches 0; Indels 9; Gaps 1;  
OY 4 KWPMPWRRKHEAPEAPIMILKK 28  
:|||||||  
Db 3 RWPMPWRRK-----IMILKK 18  
RESULT 14  
US-08-915-314-52  
Sequence 52, Application US/08915314  
Patent No. 6180604  
GENERAL INFORMATION:  
APPLICANT: Fraser, Janet R.  
APPLICANT: West, Michael H.P.  
APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
APPLICANT: Erfile, Douglas  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESS: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6180604tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 660081.405  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
FAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-915-314-52  
Query Match 36.1%; Score 87; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.3e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 LKKPMPWRRK 13  
:|||||||  
Db 1 LKKPMPWRRK 12  
RESULT 15  
US-08-915-314-74  
Sequence 74, Application US/08915314  
Patent No. 6180604  
GENERAL INFORMATION:  
APPLICANT: Fraser, Janet R.  
APPLICANT: West, Michael H.P.  
APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
APPLICANT: Erfile, Douglas  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN  
NUMBER OF SEQUENCES: 90  
CORRESPONDENCE ADDRESS:  
ADDRESS: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,314  
FILING DATE: 20-AUG-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6180604tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 660081.405  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
FAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-915-314-74  
Query Match 35.7%; Score 86; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPWPWPWR 12  
| | | | | | | | | |  
Db 1 ILKKPWPWPWR 12

Search completed: December 11, 2002, 15:39:59  
Job time : 16 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 11, 2002, 15:34:29 ; Search time 35 Seconds  
(without alignments)  
140.865 Million cell updates/sec

Title: US-09-444-281-35-27-35  
Perfect score: 241  
Sequence: 1 ILKKWPMWPMRRKHEAPEADPEPILMLKKWPMWPMRRK 37

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq.101002:\*

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141.5	58.7	28	21	AAV91800
2	137.5	57.1	27	19	AAW66363
3	133.5	55.4	63	21	AAV44668
4	133.5	55.4	63	21	AAV57142
5	99.5	41.3	16	18	AAW12899
6	99.5	41.1	21	19	AAV24582
7	99	41.1	21	21	AAV91806
8	98	40.7	21	21	AAV24571
9	98	40.7	21	21	AAV91808
10	93.5	38.8	20	19	AAV24570

11	93.5	38.8	20	21	AAV91807	Amino acid sequenc
12	91	37.8	13	18	AAW12873	Antimicrobial cati
13	91	37.8	13	19	AAV24609	Indollicidin analog
14	91	37.8	13	19	AAW66378	Cationic peptide o
15	91	37.8	13	21	AAW71690	Cationic peptide M
16	91	37.8	13	21	AAV94495	MB1-11 peptide der
17	91	37.8	13	21	AAV92795	Indollicidin analog
18	91	37.8	13	21	AAV91773	Amino acid sequenc
19	91	37.8	13	21	AAV91774	Amino acid sequenc
20	91	37.8	13	21	AAV91818	Amino acid sequenc
21	91	37.8	13	21	AAV91819	Amino acid sequenc
22	91	37.8	13	21	AAV91820	Amino acid sequenc
23	91	37.8	13	23	ABB81254	Cp11-NH2 antibacte
24	91	37.8	14	19	AAV24583	Indollicidin analog
25	91	37.8	14	21	AAV91811	Amino acid sequenc
26	89	36.9	15	18	AAW13802	Antimicrobial cati
27	87.5	36.3	21	19	AAV24552	Indollicidin analog
28	87.5	36.3	21	19	AAW66376	Cationic peptide o
29	87.5	36.3	21	21	AAV91796	Amino acid sequenc
30	87	36.1	12	19	AAV24580	Indollicidin analog
31	87	36.1	12	21	AAV91804	Amino acid sequenc
32	86	35.7	12	18	AAW12877	Antimicrobial cati
33	86	35.7	12	19	AAV24615	Indollicidin analog
34	86	35.7	12	21	AAV91833	Amino acid sequenc
35	86	35.7	13	18	AAW12896	Antimicrobial cati
36	86	35.7	13	19	AAV24613	Indollicidin analog
37	86	35.7	13	19	AAV24572	Indollicidin analog
38	86	35.7	13	21	AAV91803	Amino acid sequenc
39	86	35.7	13	21	AAV91812	Amino acid sequenc
40	86	35.7	14	19	AAV24573	Indollicidin analog
41	86	35.7	14	21	AAV91813	Amino acid sequenc
42	85.5	35.5	16	18	AAW12882	Antimicrobial cati
43	85	35.3	12	19	AAV24586	Indollicidin analog
44	85	35.3	12	21	AAV91828	Amino acid sequenc
45	85	35.3	13	18	AAW27179	Antimicrobial cati

## ALIGNMENTS

RESULT 1	AAV91800	standard; Peptide; 28 AA.
ID	AAV91800	
XX	AAV91800:	
AC	AAV91800:	
XX		
DT	06-JUN-2000 (first entry)	
XX		
DE	Amino acid sequence of cationic peptide MBI 11B20CN.	
XX		
KW	Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;	
KW	leukemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;	
KW	breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;	
KW	multidrug resistance.	
XX		
OS	Synthetic.	
XX		
PN	WO965506-A2.	
XX		
PD	23-DEC-1999.	
XX		
PF	14-JUN-1999; 99WC-CA00552.	
XX		
PR	12-JUN-1998; 98US-0096541.	
XX		
PA	(MICR-) MICROLOGIX BIOTECH INC.	
XX		
PI	Friedland HD, Krieger TJ, Taylor R, Erfile D, Fraser JR, West MHP;	
XX	WPI; 2000-223549/19.	
DR		
XX	Novel pharmaceutical composition containing optionally activated	
PT	polyoxalkylene-modified cationic peptides, useful for treating tumours	

PT  
XX  
PS  
XX  
XX  
Claim 1; Page 15; 94pp; English.

CC This sequence represents a cationic peptide amino acid sequence, which  
CC can be used in the pharmaceutical composition of the invention. The  
CC invention relates to a pharmaceutical composition containing at least one  
CC activated polyoxalkylene (APO)-modified cationic peptide. The  
CC modification of peptides with APO increases their activity against tumour  
CC cells, including those with a multidrug resistant phenotype. The  
CC pharmaceutical composition can be used to treat tumours, specifically  
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,  
CC cervix, uterus, skin, prostate, liver and colon.

SQ Sequence 28 AA;

Query Match 58.7%; Score 141.5; DB 21; Length 28;  
Best Local Similarity 58.8%; Pred. No. 3.6e-11;  
Matches 20; Conservative 2; Mismatches 1; Indels 11; Gaps 1;

QY 4 KMPWMPWRRKHEAPEPEPIMILKKMPWMPWRRK 37  
:||||||| | :|||||||  
Db 3 RMPWMPWRRK-----ILMRMPWMPWRRK 25

RESULT 2  
AAM66363  
ID AAM66363 standard; peptide; 27 AA.  
XX AAM66363;  
XX  
-DT 12-JAN-1999 (first entry)  
XX  
DE Indolicidin analogue MBI 11B20.  
XX  
XX Indolicidin analogue; resistance; cationic peptide; antibiotic;  
XX bacterial infection; tolerance; antibacterial; microorganism;  
XX bacteria; fungus; parasite; virus.  
XX  
OS Bos taurus.  
OS Synthetic.  
XX  
XX WO9840401-A2.  
XX  
XX 17-SEP-1998.  
XX  
XX 10-MAR-1998; 98WO-CA00190.  
XX  
XX 25-FEB-1998; 98US-0030619.  
XX 10-MAR-1997; 97US-0040649.  
XX 20-AUG-1997; 97US-0915314.  
XX 26-SEP-1997; 97US-0060099.  
XX  
XX (MICR-) MICROLOGIX BIOTECH INC.  
XX  
XX Fraser JR, McNICOL PJ, West MHP;  
XX  
XX WPI; 1998-520800/44.  
XX  
XX  
XX New indolicidin peptide analogues - useful for, e.g. enhancing  
XX activity of antibiotic or overcoming tolerance, acquired resistance  
XX or inherent resistance of microorganisms  
XX  
XX  
XX Claim 1; Page 91; 105pp; English.

CC The present sequence represents an indolicidin analogue. The present  
CC invention describes compositions and methods for treating infection,  
CC especially bacterial infections. The compositions and methods use  
CC cationic peptides in combination with an antibiotic agent which are  
CC then administered to a patient to enhance the activity of the antibiotic  
CC agent, to overcome: (a) tolerance; (b) acquired resistance; and (c)  
CC inherent resistance. The combinations of antibiotics and cationic  
CC peptides can provide synergistic activity against a microorganism that

CC is tolerant, inherently resistant, or has acquired resistance to an  
CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,  
CC parasites and viruses.

SQ Sequence 27 AA;

Query Match 57.1%; Score 137.5; DB 19; Length 27;  
Best Local Similarity 61.8%; Pred. No. 1.1e-10;  
Matches 21; Conservative 2; Mismatches 0; Indels 11; Gaps 2;

QY 4 KMPWMPWRRKHEAPEPEPIMILKKMPWMPWRRK 37  
:||||||| | :|||||||  
Db 3 RMPWMPWRRK-----ML-RMPWMPWRRK 25

RESULT 3  
AAV44668  
ID AAV44668 standard; Protein; 63 AA.  
XX  
XX AAV44668;  
XX  
-DT 18-APR-2000 (first entry)  
XX  
XX  
DE Poly-(Indol (1-13)-Met-Ala-Arg-Ile-Ala-Met)3 protein.  
XX  
XX Crosslinked indolicidin analog; X-indolicidin; poly-Indol 1-13;  
XX stability; bovine neutrophil; antimicrobial; antibacterial; fungicide;  
XX prozoocicide; vitruicide; anti-HIV; human immunodeficiency virus-1;  
XX HIV-1; gram positive bacteria; gram negative; Staphylococcus aureus;  
XX Escherichia coli; Salmonella typhimurium; yeast; fungi; protozoa;  
XX Candida albicans; Cryptococcus neoformans; Giardia; Acanthamoeba;  
XX hexapeptide spacer.  
XX  
XX Synthetic.  
XX Bos sp.  
XX  
XX  
OS  
OS  
FH Key Location/Qualifiers  
FH Region 1..5  
FT Cleavage-site /label= Enterokinase\_recognition\_site  
FT 5..6  
FT Cleavage-site /label= Enterokinase\_cleavage\_site  
FT 6..7  
FT Cleavage-site /label= Cyanogen\_bromide\_cleavage\_site  
FT 20..21  
FT Cleavage-site /label= Cyanogen\_bromide\_cleavage\_site  
FT 25..26  
FT Cleavage-site /label= Cyanogen\_bromide\_cleavage\_site  
FT 39..40  
FT Cleavage-site /label= Cyanogen\_bromide\_cleavage\_site  
FT 44..45  
FT Cleavage-site /label= Cyanogen\_bromide\_cleavage\_site  
FT 58..59  
FT Cleavage-site /label= Cyanogen\_bromide\_cleavage\_site  
FT 20..25  
FT Region /label= Hexapeptide\_spacer  
FT Region 39..44  
FT Region /label= Hexapeptide\_spacer  
FT 58..63  
FT Region /label= Hexapeptide\_spacer

WO965510-A1.  
23-DEC-1999.  
20-MAY-1999; 99WO-US11165.  
18-JUN-1998; 98US-0099631.  
(RESC ) UNIV CALIFORNIA.  
Selsled ME, Osapay K;  
WPI; 2000-147133/13.

```

DR      N-PSDB:AAZ49764.
XX      Crosslinked indolicidin analogs with antimicrobial activity against
PT      bacteria, yeast, fungi, protozoa and viruses
XX
XX      Example 1C; Fig 1; 53pp; English.
XX
CC      The patent discloses crosslinked analogs of indolicidin (Indol 1-13)
CC      which is a naturally occurring peptide isolated from bovine neutrophils
CC      and has antimicrobial activity. The crosslinked indolicidin
CC      (X-indolicidin) analogs are stable and have antimicrobial activity
CC      against gram positive and negative bacteria (e.g. Staphylococcus aureus,
CC      Escherichia coli and Salmonella typhimurium), yeasts and fungi (e.g.
CC      Candida albicans, Cryptococcus neoformans), protozoa (e.g. Giardia
CC      species and Acanthamoeba species), and viruses (e.g. HIV-1).
CC      They can be used for reducing or inhibiting the growth or survival of
CC      microorganisms in an environment e.g. a food or food product, a
CC      solution, an inanimate object comprising a surface, or a mammal.
CC      The present sequence is a protein comprising three
CC      copies of Indol 1-13 each separated by a hexapeptide spacer sequence.
CC      A recombinant construct encoding this sequence was used for the
CC      expression of Indol-homoserine (Hse) analog. The ability of
CC      Indol-Hse analog to maintain antimicrobial activity provides a means to
CC      produce X-indolicidin analog precursors in sufficient quantities.
XX
XX      Sequence 63 AA:
XX
XX      Query Match          55.4%; Score 133.5; DB 21; Length 63;
XX      Best Local Similarity 64.7%; Pred. NO. 9e-10;
XX      Matches 22; Conservative 0; Mismatches 5; Indels 7; Gaps 2
XX
OY      4 KWPMPWRRKHEAPEPEIMLK-KWPMPWRR 36
XX      |||||
XX      11 KWPMPWRRM-----ARIAMILPKWPMPWRR 38
XX
XX      RESULT 4
XX      AAY57142
XX      ID AAY57142 standard; Protein; 63 AA.
XX      AC AAY57142;
XX      DT 28-FEB-2000 (first entry)
XX
XX      Indolicidin fusion peptide amino acid sequence.
XX
XX      Indolicidin analogue; antimicrobial activity; helminth: bacteria; virus;
XX      treatment; inhibit growth; micro-organism; contact lens solution;
XX      transgenic plant; surgical instrument; yeast; fungi; protozoa.
XX      Synthetic.
XX      OS
XX      PN WO9558141-A1.
XX      PD 18-NOV-1999.
XX      PF 05-MAY-1999; 99WO-US09942.
XX      PR 12-MAY-1998; 98US-0076227.
XX      PA (RESC ) UNITV CALIFORNIA.
XX      PI
XX      PI
XX      DR MPI: 2000-053028/04.
XX      DR N-PSDB:AAZ45123.
XX
XX      New indolicidin analogues, active against bacteria, yeast, fungi,
XX      protozoa and virus, used for, e.g. treating infections -
XX      Disclosure; Fig 6; 62pp; English.
XX
XX      This is the amino acid sequence of an example of a fusion protein which

```

CC	consists of an indolicidin analogue linked to another peptide.							
CC	Peptides AAY7109-Y57138 and AAY57143-Y57144 are new indolicidin							
CC	analogues, which have a homoserine residue and/or a truncated amino							
CC	terminal region. The analogues have the following amino acid sequence:							
CC	Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-Xaa6-Pro-Xaa6-Xaa7-Xaa7'-Xaa8							
CC	Where:							
CC	Xaa1 = Ile, Leu, Val, Ala, Gly or absent;							
CC	Xaa2 = Ile, Leu, Val, Ala, Gly or absent;							
CC	Xaa3 = Pro or absent;							
CC	Xaa4 = Trp, Phe or absent;							
CC	Xaa5 = Arg, Lys or absent;							
CC	Xaa6 = Trp or Phe;							
CC	Xaa7 = Arg, Lys or absent;							
CC	Xaa8 = homoserine (Hse), Met, Met-Xaa9-Met or absent, and							
CC	Xaa9 = at least one amino acid;							
CC	provided that if Xaa1 is present, Xaa8 = Hse, Met or Met-Xaa9-Met;							
CC	and further provided that: if Xaa2 is absent, Xaa1 is absent; if Xaa3 is							
CC	absent, Xaa1 and Xaa2 are absent; if Xaa4 is absent, Xaa1, Xaa2 and Xaa3							
CC	are absent; and if Xaa5 is absent, Xaa1, Xaa2, Xaa3 and Xaa4 are absent.							
CC	The indolicidin analogues can be used to create a fusion polypeptide							
CC	consisting of the analogue linked to a peptide. The indolicidin							
CC	analogues have antimicrobial activity against gram positive bacteria,							
CC	gram negative bacteria, yeast, fungus, protozoa and viruses (e.g. HIV-1).							
CC	They are also active against helminths. The analogues can be used for							
CC	reducing or inhibiting growth or survival of a microorganism. They can be							
CC	used for treating infections. They can also be included in a liquid such							
CC	as water or an aqueous solution, e.g. contact lens solution. The							
CC	analogues have potential uses in food products, and in objects such as							
CC	the surface of an instrument used to prepare food or to perform surgery.							
CC	Transgenic plants or animals useful in the food industry can be produced							
CC	by introducing a nucleic acid molecule encoding an indolicidin analogue							
CC	into the germ-line cells of such organisms.							
XX	SQ	Sequence	63 AA;					
OY	Query Match	55.4%;	Score 133.5;	DB 21;	Length 63;			
	Best Local Similarity	64.7%;	Pred. No. 9e-10;					
	Matches 22;	Conservative 0;	Mismatches 5;	Indels 7;	Gaps 2;			
Dd	4 KMPWMPRRKHEAPEEPIIMLK-KMPWMPRR 36       1 KMPWMPRRM-----ARIAMILPKWKVMPWR 38							
RESULT 5								
AAM12899								
ID	AAM12899 standard: peptide; 16 AA.							
XX	AAM12899;							
XX	DT	10-DEC-1997 (first entry)						
DE	Antimicrobial cationic peptide CP-26.							
XX								
KW	Bacterial; viral; antitumour; food; preservative; inhibitor; growth;							
KW	bacterium; yeast; endotoxaemia; sepsis; antibiotic; fungal;							
KW	antiviral; Candida albicans; steriliant; Salmonella; Yersina;							
KW	Shigella.							
OS	Synthetic.							
XX	PN	WO9708199-A2.						
PD	06-MAR-1997.							
PF	23-AUG-1996;	96WO-IB00996.						
PR	23-AUG-1995;	95US-0002687.						
PA	(UYBR-) UNIV BRITISH COLUMBIA.							
PI	Falla TJ, Gough M, Hancock RW;							

DR WPI: 1997-179179/16.

XX Cationic peptide(s) having anti-microbial activity - used for the  
 PT inhibition of bacterial and viral growth, as an antitumour agent,  
 PT and as a food preservative  
 XX  
 PS Claim 3; Page 66; 89pp; English.

XX The present sequence represents a specifically claimed novel isolated  
 CC cationic peptide which has antimicrobial activity. The amino acid  
 CC sequence of antimicrobial cationic peptides (including the present  
 CC sequence) is selected from: X1X1ProX2X3X2Pro(X2X2Pro)X2X3(X5)O;  
 CC X1X1ProX2X3X4(X5)ProX2X3X3; X1X1X3(ProTrp)uX3X2X5X2X5X2(X5)O;  
 CC X1X1X3X3X2Pro(X2X2Pro)X2(X5)m: where m = 1-5; n = 1-2; o = 2-5; r  
 CC = 0-8; u = 0-1; X1 = Ile, Leu, Val, Phe, Tyr, Trp or Met; X2 = Trp or  
 CC Phe; X3 = Arg or Lys; X4 = Trp or Lys; and X5 = Phe, Trp, Arg, Lys or  
 CC Pro. The peptides are preferably amidated or carboxymethylated. The  
 CC peptides may be used in methods for inhibiting the growth of a bacterium  
 CC or yeast, or for inhibiting an endotoxaemia or sepsis associated  
 CC disorder in a subject. The peptides have a broad activity against  
 CC antibiotic resistant bacteria, combined with activity against the  
 CC medically important fungus *Candida albicans*. In addition, the peptides  
 CC are useful as antitumour agents and/or antiviral agents. The peptides  
 CC may be used as sterilants or preservatives of materials susceptible to  
 CC microbial or viral contamination, e.g. in processed foods to inhibit  
 CC *Salmonella*, *Yersinia* and *Shigella*. The peptides are compact and tend to  
 CC have a unique polypyrrolone type II extended helix structure that permits  
 CC them to span the membrane with relatively few amino acids. The peptides  
 CC possess the ability to work synergistically with antibiotics, and in  
 CC addition, some of them possess anti-endotoxin activity.

XX  
 SQ Sequence 16 AA:

Query Match 41.3%; Score 99.5; DB 18; Length 16;  
 Best Local Similarity 43.2%; Pred. No. 3.7e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

OY 1 ILKKWPMWRKHEAPEPIMILKKWPMWRK 37  
 |||||  
 DB 1 ILKKWPMWRKHEAPEPIMILKKWPMWRK 16

RESULT 6  
 AAY24582  
 ID AAY24582 standard; peptide; 21 AA.  
 AC AAY24582;  
 XX  
 DT 18-AUG-1999 (first entry)  
 XX  
 DE Indolicidin analogue #34.  
 XX  
 KW Indolicidin; bacterial infection; photo-oxidised solubiliser;  
 KW antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;  
 KW additive; shampoo; soap; insecticide; herbicide; preservative;  
 KW food; technical material.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9807745-A2.  
 XX  
 PD 26-FEB-1998.  
 XX  
 PF 21-AUG-1997; 97WO-US14779.  
 XX  
 PR 13-JAN-1997; 97US-0034949.  
 PR 21-AUG-1996; 96US-0024754.  
 XX  
 PA (MICR-) MICROLOGIX BIOTECH INC.  
 XX  
 PI Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;  
 XX  
 DR WPI: 1998-169090/15.

XX New indolicidin analogues with antimicrobial activity and related  
 PT nucleic acid - vectors, transformed cells and antibodies, also  
 PT conjugates with polyoxyalkylene glycol and fatty acid to reduce  
 PT toxicity, useful therapeutically, as disinfectants etc.  
 XX  
 PS Claim 13; Page 89; 129pp; English.

XX AAY24549 to AAY24615 represent indolicidin analogues of formulae  
 CC (I)-(VIII) containing up to 25 amino acids (aa): PX2XXZXB (I), BX2XXZXB  
 CC (II), BBX2XXZXB (III), BZX2XXZBBn(A)nmlBBBGS (IV), BX2XXZBB(A)nM  
 CC (V), LBN2XXZnXnXnR (VI), LK2XXZXXRR (VII) and BBX2XXZXB (VIII).  
 CC Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa;  
 CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;  
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat  
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);  
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or  
 CC trematodes) or viruses. Typical of very many pathogens that can be  
 CC controlled are *Leishmania*, *Trypanosoma*, *Ascaris lumbricoides*, *Fasciola*  
 CC *hepatica*, *Klebsiella pneumoniae*, *Bordetella pertussis*, *Staphylococcus*  
 CC *aureus*, *Listeria*, *Clostridium*, rotavirus and papilloma virus. Compounds  
 CC derived from the analogues may be used similarly: the compounds may  
 CC also be prepared from antibiotics or antiarrhythmic agents. The analogues  
 CC may be used therapeutically or to coat medical devices; also they are  
 CC useful as surface disinfectants, as additives to shampoo or soaps, as  
 CC insecticides or herbicides, or as preservatives for foods and technical  
 CC materials. The analogues are administered by injection, lavage, orally  
 CC or topically, generally at 0.1-50 mg/Kg. These analogues have a broader  
 CC spectrum of activity than indolicidin and modification as compounds  
 CC reduces their toxicity.

XX  
 SQ Sequence 21 AA:

Query Match 41.1%; Score 99; DB 19; Length 21;  
 Best Local Similarity 64.3%; Pred. No. 5.9e-06;  
 Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

OY 1 ILKKWPMWRKHEAPEPIMILKK 28  
 |||||  
 DB 1 ILKKWPMWRKHEAPEPIMILKK 18

RESULT 7  
 AAY91806  
 ID AAY91806 standard; peptide; 21 AA.  
 AC AAY91806;  
 XX  
 DT 06-JUN-2000 (first entry)  
 XX  
 DE Amino acid sequence of cationic peptide MBI 11D4CN.  
 XX  
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;  
 KW leukaemia; polyoxyalkylene-modified; Apo; lymphoma; multiple myeloma;  
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;  
 KW multidrug resistance.  
 XX  
 OS Synthetic.  
 XX  
 PN WO965506-A2.  
 XX  
 PD 23-DEC-1999.  
 XX  
 PF 14-JUN-1999; 99WO-CA00552.  
 XX  
 PR 12-JUN-1998; 98US-0096541.  
 XX  
 PA (MICR-) MICROLOGIX BIOTECH INC.  
 XX  
 PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;  
 XX  
 DR WPI: 2000-223549/19.



PT Novel pharmaceutical composition containing optionally activated  
PT polyoxalkylene-modified cationic peptides, useful for treating tumours  
PS  
PS Disclosure: Page 15; 94pp; English.  
XX  
CC This sequence represents a cationic peptide amino acid sequence, which  
CC can be used in the pharmaceutical composition of the invention. The  
CC invention relates to a pharmaceutical composition containing at least one  
CC activated polyoxalkylene (APO)-modified cationic peptide. The  
CC modification of peptides with APO increases their activity against tumour  
CC cells, including those with a multidrug resistant phenotype. The  
CC pharmaceutical composition can be used to treat tumours, specifically  
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,  
CC cervix, uterus, skin, prostate, liver and colon.  
SQ Sequence 21 AA;  
SQ  
Query Match 41.1%; Score 99; DB 21; Length 21;  
Best Local Similarity 64.3%; Pred. No. 5.9e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 1 ILKKPMWPMRRKHAEPEAEPIMLKK 28  
Db 1 ILKKPMWPMRRK-----IMILKK 18  
RESULT 8  
AAAY24571  
ID AAY24571 standard; peptide; 21 AA.  
AC AAY24571;  
XX  
DT 18-AUG-1999 (first entry)  
XX  
DE Indolicidin analogue #23.  
XX  
KM Indolicidin: bacterial infection; photo-oxidised solubiliser;  
KM antimicrobial; antibiotic; antitryptinase; surface disinfectant;  
KM additive; shampoo; soap; insecticide; herbicide; preservative;  
KM food; technical material.  
XX  
OS Synthetic.  
XX  
PN WO9807745-A2.  
XX  
PD 26-FEB-1998.  
XX  
PF 21-AUG-1997; 97WO-US14779.  
XX  
PR 13-JAN-1997; 97US-0034949.  
PR 21-AUG-1996; 96US-0024754.  
XX  
PA (MICR-) MICROLOGIX BIOTECH INC.  
XX  
PI Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;  
PI WPI; 1998-169090/15.  
DR  
XX  
XX New indolicidin analogues with antimicrobial activity and related  
PT nucleic acid - vectors, transformed cells and antibodies, also  
PT conjugates with polyoxalkylene glycol and fatty acid to reduce  
PT toxicity, useful therapeutically, as disinfectants etc.  
PS  
XX Claim 12; Page 89; 129pp; English.  
XX  
CC AAY24549 to AAY24615 represent indolicidin analogues of formulae  
CC (I)-(VIII) containing up to 25 amino acids (aa): RxxXxxXB (I), BxxXxxXB  
CC (II), BxxXxxXxxB (III), BxxXxxBxxB(AA)nMILBAGS (IV), BxxXxxB(AA)nM  
CC (V), LBnXxxXxxXxxR (VI), LKxxXxxXxxR (VII) and BxxXxxXxxBB (VIII).  
CC Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,  
CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;  
CC in (VIII) at least 2 X = F or Y. The analogues are used to treat

CC infections caused by bacteria (Gram positive or negative, or anaerobic);  
CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or  
CC trematodes) or viruses. Typical of very many pathogens that can be  
CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola  
CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus  
CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds  
CC derived from the analogues may be used similarly; the compounds may  
CC also be prepared from antibiotics or antitryptinase agents. The analogues  
CC may be used therapeutically or to coat medical devices; also they are  
CC useful as surface disinfectants, as additives to shampoo or soaps, as  
CC insecticides or herbicides, or as preservatives for foods and technical  
CC materials. The analogues are administered by injection, lavage, orally  
CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader  
CC spectrum of activity than indolicidin and modification as compounds  
CC reduces their toxicity.  
SQ Sequence 21 AA;  
SQ  
Query Match 40.7%; Score 98; DB 19; Length 21;  
Best Local Similarity 64.3%; Pred. No. 7.8e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;  
QY 1 ILKKPMWPMRRKHAEPEAEPIMLKK 28  
Db 1 ILKKPMWPMRRK-----IMILKK 18  
RESULT 9  
AAAY91808  
ID AAY91808 standard; Peptide; 21 AA.  
AC AAY91808;  
XX  
DT 06-JUN-2000 (first entry)  
XX  
DE Amino acid sequence of cationic peptide MBI 11D6CN.  
XX  
KM Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;  
KM leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;  
KM breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;  
KM multidrug resistance.  
XX  
OS Synthetic.  
XX  
PN WO965506-A2.  
XX  
PD 23-DEC-1999.  
XX  
PF 14-JUN-1999; 99WO-CA00552.  
XX  
PR 12-JUN-1998; 98US-0096541.  
XX  
PA (MICR-) MICROLOGIX BIOTECH INC.  
XX  
PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;  
PI WPI; 2000-223549/19.  
DR  
XX  
XX Novel pharmaceutical composition containing optionally activated  
PT polyoxalkylene-modified cationic peptides, useful for treating tumours  
PT  
PS  
XX Disclosure: Page 15; 94pp; English.  
XX  
CC This sequence represents a cationic peptide amino acid sequence, which  
CC can be used in the pharmaceutical composition of the invention. The  
CC invention relates to a pharmaceutical composition containing at least one  
CC activated polyoxalkylene (APO)-modified cationic peptide. The  
CC modification of peptides with APO increases their activity against tumour  
CC cells, including those with a multidrug resistant phenotype. The  
CC pharmaceutical composition can be used to treat tumours, specifically  
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,  
CC cervix, uterus, skin, prostate, liver and colon.

XX Sequence 21 AA; Score 98; DB 21; Length 21;  
 SQ Query Match 40.7%; Pred. No. 7.8e-06;  
 Best Local Similarity 64.3%; Mismatches 0; Indels 10; Gaps 1;  
 Matches 18; Conservative 0; Indels 10; Gaps 1;

OY 1 ILKKPMPWRRKHEAPEAPIMILKK 28  
 |||||  
 DB 1 ILKKPMPWRRR-----MILKK 18

RESULT 10  
 AAY24570  
 ID AAY24570 standard; peptide; 20 AA.  
 XX  
 AC AAY24570;  
 XX  
 DT 18-AUG-1999 (first entry)  
 XX  
 DE Indolicidin analogue #22.  
 XX  
 KM Indolicidin; bacterial infection; photo-oxidised solubilisier;  
 KM antimicrobial; antibiotic; antiaertrymic; surface disinfectant;  
 KM additive; shampoo; soap; insecticide; herbicide; preservative;  
 KM food; technical material.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9807745-A2.  
 XX  
 PD 26-FEB-1998.  
 XX  
 PF 21-AUG-1997; 97WO-US14779.  
 XX  
 PR 13-JAN-1997; 97US-0034949.  
 XX  
 PR 21-AUG-1996; 96US-0024754.  
 XX  
 PA (MICR-) MICROLOGIX BIOTECH INC.  
 XX  
 PI Erifle D. Fraser JR, Krieger TJ, Taylor R, West MH;  
 XX  
 DR WPI; 1998-169090/15.  
 XX  
 PT New indolicidin analogues with antimicrobial activity and related  
 PT nucleic acid- vectors, transformed cells and antipodes, also  
 PT conjugates with polyoxyalkylene glycol and fatty acid to reduce  
 PT toxicity, useful therapeutically, as disinfectants etc.  
 XX  
 PS Claim 12; Page 89; 129pp; English.  
 XX

AAY24549 to AAY24615 represent indolicidin analogues of formulae  
 CC (I)-(VII) containing up to 25 amino acids (aa): RXXXXXB (I), BXXXXXB  
 CC (II), BXXXXXB (III), BXXXXBBn(AA)nMILBAGS (IV), BXXXXXB(AA)nM  
 CC (V), LBnXXnXXnXXnXXR (VI), LKnXXnXXnXXR (VII) and BXXXXXB (VIII).  
 CC Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,  
 CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;  
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat  
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);  
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or  
 CC trematodes) or viruses. Typical of very many pathogens that can be  
 CC controlled are leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola  
 CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus  
 CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds  
 CC derived from the analogues may be used similarly; the compounds may  
 CC also be prepared from antibiotics or antiaertrymic agents. The analogues  
 CC may be used therapeutically or to coat medical devices; also they are  
 CC useful as surface disinfectants, as additives to shampoo or soaps, as  
 CC insecticides or herbicides, or as preservatives for foods and technical  
 CC materials. The analogues are administered by injection, lavage, orally  
 CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader  
 CC spectrum of activity than indolicidin and modification as compounds  
 CC reduces their toxicity.

XX Sequence 20 AA; Score 93.5; DB 19; Length 20;  
 SQ Query Match 38.8%; Pred. No. 2.7e-05;  
 Best Local Similarity 60.7%; Mismatches 0; Indels 11; Gaps 1;  
 Matches 17; Conservative 0; Indels 11; Gaps 1;

OY 1 ILKKPMPWRRKHEAPEAPIMILKK 28  
 |||||  
 DB 1 ILKKPMPWRRR-----MILKK 17

RESULT 11  
 AAY91807  
 ID AAY91807 standard; peptide; 20 AA.  
 XX  
 AC AAY91807;  
 XX  
 DT 06-JUN-2000 (first entry)  
 XX  
 DE Amino acid sequence of cationic peptide MBI 11D5CN.  
 XX  
 KM Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;  
 KM leukaemia; polyoxyalkylene-modified; Apo; lymphoma; multiple myeloma;  
 KM breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;  
 KM multidrug resistance.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9965506-A2.  
 XX  
 PD 23-DEC-1999.  
 XX  
 PF 14-JUN-1999; 99WO-CA00552.  
 XX  
 PR 12-JUN-1998; 98US-0096541.  
 XX  
 PA (MICR-) MICROLOGIX BIOTECH INC.  
 XX  
 PI Friedland HD, Krieger TJ, Taylor R, Erifle D, Fraser JR, West MHP;  
 XX  
 DR WPI; 2000-223549/19.  
 XX  
 PT Novel pharmaceutical composition containing optionally activated  
 PT polyoxyalkylene-modified cationic peptides, useful for treating tumours  
 PT  
 PS Disclosure; Page 15; 94pp; English.  
 XX

This sequence represents a cationic peptide amino acid sequence, which  
 CC can be used in the pharmaceutical composition of the invention. The  
 CC invention relates to a pharmaceutical composition containing at least one  
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The  
 CC modification of peptides with APO increases their activity against tumour  
 CC cells, including those with a multidrug resistant phenotype. The  
 CC pharmaceutical composition can be used to treat tumours, specifically  
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,  
 CC cervix, uterus, skin, prostate, liver and colon.  
 CC

RESULT 12  
 AAW12873  
 ID AAW12873 standard; peptide; 13 AA.



```

ID AAW6378 standard; peptide: 13 AA.
XX
AC AAW6378;
XX
DT 12-JAN-1999 (first entry)
XX
DE Cationic peptide of claim 15 #5.
XX
KW Indolicidin analogue; resistance; cationic peptide; antibiotic;
KW bacterial infection; tolerance; antibacterial; microorganism;
KW bacteria; fungus; parasite; virus.
XX
OS Synthetic.
XX
PN WO9840401-A2.
XX
PD 17-SEP-1998.
XX
PF 10-MAR-1998; 98WO-CA00190.
XX
PR 25-FEB-1998; 98US-0030619.
PR 10-MAR-1997; 97US-0040649.
PR 20-AUG-1997; 97US-0915314.
PR 26-SEP-1997; 97US-0060099.
XX
PA (MICR-) MICROLOGIX BIOTECH INC.
XX
PI Fraser JR, McNICOL PJ, West MHP;
XX
DR WPI; 1998-520800/44.
XX
PT New indolicidin peptide analogues - useful for, e.g. enhancing
PT activity of antibiotic or overcoming tolerance; acquired resistance
PT or inherent resistance of microorganisms
XX
PS Claim 15; Page 93; 105pp; English.
XX
'CC The present sequence represents a specifically claimed cationic peptide
CC from the present invention. The present invention describes compositions
CC and methods for treating infection, especially bacterial infections. The
CC compositions and methods use cationic peptides in combination with an
CC antibiotic agent which are then administered to a patient to enhance the
CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
CC acquired resistance; and (c) inherent resistance. The combinations of
CC antibiotics and cationic peptides can provide synergistic activity
CC against a microorganism that is tolerant, inherently resistant, or has
CC acquired resistance to an antibiotic agent. They can be used for killing
CC e.g. bacteria, fungi, parasites and viruses.
XX
SQ Sequence 13 AA:

Query Match 37.8%; Score 91; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPWPRRK 13
   |||||
Db 1 ILKKPMPWPRRK 13

RESULT 15
AAW71690
ID AAW71690 standard; Peptide: 13 AA.
XX
AC AAW71690;
XX
DT 11-JAN-1999 (first entry)
XX
DE Cationic peptide MB111 (MW 1879).
XX
KW MB111; cationic peptide; plasmid PKL1; small cryptic plasmid;
KW replication; RepA; vector; RAMP.
XX

```

```

OS Synthetic.
XX
PN WO9841636-A2.
XX
PD 24-SEP-1998.
XX
PF 16-MAR-1998; 98WO-CA00214.
XX
PR 14-MAR-1997; 97US-0040722.
XX
PA (BURI/) BURIAN J.
PA (KAYW/) KAY W W.
XX
PI Burian J, Kay WW;
XX
DR WPI; 1998-531571/45.
XX
PT Increasing plasmid copy number in a cell with the repA gene product
PT - and an small cryptic plasmid ori sequence, useful for high level
PT expression of e.g. cytokines, antigens or therapeutic proteins
XX
PS Example 13; Page 54; 82pp; English.
XX
CC MB111 is a small (mol.wt. 1879) cationic peptide. DNA encoding
CC MB111 has been incorporated into vector pR2h-B1, in which the
CC replication leader (R21) sequence of RepA (see also AAW71686) is
CC joined to 2 Hpro peptides (see also AAW71692), to provide a
CC vector for expression of MB111 in host cells. The invention
CC provides controlled replication plasmid vectors (RAMP vectors)
CC comprising a replicated origin of a small cryptic plasmid and a
CC gene encoding RepA. The vectors can reach very high levels of
CC plasmid replication, but are not lethal to the host cell, and can
CC be used to direct the high level expression of e.g. cytokines,
CC antigens and therapeutic proteins.
XX
SQ Sequence 13 AA:

Query Match 37.8%; Score 91; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPWPRRK 13
   |||||
Db 1 ILKKPMPWPRRK 13

Search completed: December 11, 2002, 15:38:44
Job time : 35 secs

```